Raytheon

SURFACE TYPE

VISIBLE/INFRARED IMAGER/RADIOMETER SUITE ALGORITHM THEORETICAL BASIS DOCUMENT

Version 4: May 2001

Eric C. Brown de Colstoun
Shawn W. Miller
Wenli Yang

Phase I Science Team (University of Maryland)
Ruth DeFries
Matt Hansen
John Townshend

RAYTHEON SYSTEMS COMPANY
Information Technology and Scientific Services
4400 Forbes Boulevard
Lanham, MD 20706

SBRS Document #: Y2402

EDR: SURFACE TYPE

Doc No: Y2402

Version: 4

Revision: 0

	Function	Name	Signature	Date
Prepared by	EDR Developer	E. C. BROWN DE COLSTOUN		
Approved by	Relevant IPT Lead	S. MILLER		
Approved by	Chief Scientist	S. MILLER		
Released by	Algorithm Lead	P. KEALY		

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF ACRONYMS

ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ATBD	Algorithm Theoretical Basis Document
AVHRR	Advanced Very High Resolution Radiometer
BOA	Bottom of Atmosphere
BOREAS	BOReal Ecosystem Atmosphere Study
BRDF	Bidirectional Reflectance Distribution Function
BT	Brightness Temperature
CV	Coefficients of Variation
DCW	Digital Chart of the World
DN	Digital Number
DT	Decision Tree
EDC	EROS Data Center
EDR	Environmental Data Record
EOS	Earth Observing System
EROS	Earth Resources Observation Systems
ETM+	Enhanced Thematic Mapper Plus
FAO	Food and Agriculture Organization
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FIFE	First ISLSCP Field Experiment
GCM	General Circulation Model
GLCTS	Global Land Cover Test Sites
HAPEX	Hydrologic and Atmospheric Pilot Experiment
HCS	Horizontal Cell Size
ICSR	Issue/Clarification/Support Request
IGBP	International Geosphere-Biosphere Programme
IGBP-DIS	IGBP-Data and Information Systems
IPO	Integrated Program Office
ISLSCP	International Satellite Land Surface Climatology Project
JHU	Johns Hopkins University
LAI	Leaf Area Index
LCWG	IGBP-Land Cover Working Group
LQF	Land Quality Flag
LSP	Land Surface Parameterization
LST	Land Surface Temperature
LVQ	Learning Vector Quantization
MISR	Multi-angle Imaging Spectroradiometer
MLC	Maximum Likelihood Classifier
MODIS	Moderate Resolution Imaging Spectroradiometer

MODLAND	MODIS Land Discipline Team
MRLC	Multi-Resolution Land Characteristics
MSS	Multi-Spectral Scanner
MTF	Modulation Transfer Function
NALC	North America Land Characterization Project
NASA	National Aeronautics and Space Administration
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NedL	Noise equivalent Delta Radiance
NN	Neural Net
NOAA	National Oceanic and Atmospheric Administration
NPOESS	National Polar-orbiting Operational Environmental Satellite System
NPP	Net Primary Productivity
OLVQ	Optimized-Learning-Rate LVQ
PAL	Pathfinder AVHRR Land
PDR	Preliminary Design Review
POLDER	Polarization and Directionality of the Earth's Reflectances
SAVI	Soil Adjusted Vegetation Index
SBRS	Santa Barbara Remote Sensing
SDR	Sensor Data Record
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SOC	Satellite Operations Command
SPOT	Système Probatoire pour L'Observation de la Terre
SRD	Sensor Requirements Document
SRR	System Requirement Review
STATSGO	State Soil Geographic Database
SZA	Solar Zenith Angle
TBD	To Be Decided
TBR	To Be Resolved
TM	Thematic Mapper
TOA	Top of Atmosphere
UMD	University of Maryland
USDA	U.S. Department of Agriculture
VIIRS	Visible/Infrared Image Radiometer Suite
VWG	LCWG-Validation Working Group
VZA	View Zenith Angle

ABSTRACT

Accurate information about the actual distribution of land surface types on the Earth's surface is critical to many studies of global change. This Algorithm Theoretical Basis Document (ATBD) describes the approach taken to retrieve 17 land cover types specified by the Integrated Program Office (IPO) from data acquired by the Visible/Infrared Imager/Radiometer Suite (VIIRS) on an operational basis. This instrument is to be put in orbit early in the 21st century as a part of the National Polar-orbiting Environmental Satellite System (NPOESS).

The Surface Type Environmental Data Record (EDR) will be produced at the highest spatial resolution common to the VIIRS bands used (approximately 1 km). The operational EDR will be developed around a global 1 km VIIRS Quarterly Surface Types Intermediate Product (IP) which will be produced every three months from the accumulation of the previous 12 months of VIIRS data. During each successive three month period, the VIIRS Quarterly Surface Types IP will be re-delivered for every VIIRS orbit in conjunction with the current VIIRS Vegetation Index, Snow Cover/Depth, and Active Fires EDRs. The fraction of green vegetation cover present per cell will also be provided as a part of this EDR, further accommodating users who may require instantaneous information about the surface conditions associated with each surface type.

The VIIRS Quarterly Surface Types algorithm will be run in a supervised classification mode, using global training data specifically tailored to the IPO surface types, and temporal metrics developed from 12 months of VIIRS visible, infrared and thermal spectral band information. Training data will be screened through cross-validation using a decision tree classifier, and used to classify the remaining cells using the same decision tree classifier. A Quarterly Continuous Fields IP tailored after the Moderate Resolution Imaging Spectroradiometer (MODIS) continuous fields products will also be provided every three months using the same training data and metrics used in the production of the Quarterly Surface Types IP. Two additional quarterly products will be provided to assist the varied requirements of several other VIIRS EDRs, namely the VIIRS Surface Types-Biomes and the VIIRS Surface Types-Olson Intermediate Products (IPs). These will result from the aggregation of the Quarterly Surface Types IP into the appropriate surface types.

Before launch, the algorithm has been, and will be, extensively tested and validated using several data sets. These include the Pathfinder AVHRR Land (PAL) global data set at 8 km resolution, the 1 km global land cover classification from the Earth Resources Observation Systems (EROS) Data Center (EDC), and both the 8 km and 1 km global land cover products from the Department of Geography at the University of Maryland (UMD). After launch, the algorithm will be continually evaluated against data from the Global Land Cover Test Sites (GLCTS) as well as data and products from MODIS and its network of validation sites.

The algorithm presented in this ATBD culminates in the development of the first, fully-automated, global 1-km surface type classification available worldwide, producing typing accuracies which greatly exceed threshold performance and move well beyond those that have been reported in the literature for similar products. This further highlights the strong commitment of this team to provide a flexible, robust, and accurate Surface Type algorithm that will ultimately satisfy the requirements of a wide variety of potential users.

1.0 INTRODUCTION

1.1 PURPOSE

The accurate representation of actual terrestrial surface types from regional to global scales is an important element for many applications, including land management, implementation of national and international policies related to bio-diversity and climate change, and global change studies as input to climate, biogeochemical, and/or hydrological models. Land cover classifications at the local to regional scales have a significant heritage in the remote sensing literature, dating back to studies utilizing data from Landsat's Multi-Spectral Scanner (MSS) and Thematic Mapper (TM), and the Système Probatoire pour l'Observation de la Terre (SPOT), among others.

More recently, continental and global scale land cover classifications at spatial resolutions ranging from 1 to 20 km have been performed from data acquired by the Advanced Very High Resolution Radiometer (AVHRR), flown onboard the National Oceanic and Atmospheric Administration's NOAA-n series of satellite platforms. In the very near future, improved global land cover products will be available from the Moderate Resolution Imaging Spectroradiometer (MODIS), launched as a part of the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS).

In this Algorithm Theoretical Basis Document (ATBD) we present a general overview of the operational retrieval of global land cover information as a part of the Surface Type Environmental Data Record (EDR) and using data acquired by the future Visible/Infrared Imager/Radiometer Suite (VIIRS), one of the sensors to be included within the National Polar-orbiting Operational Environmental Satellite System (NPOESS).

1.2 SCOPE

This document is similar in structure to the ATBDs produced for the EOS sensors, particularly those of the Multi-angle Imaging Spectroradiometer (MISR) (http://eospso.gsfc.nasa.gov/atbd/misrtables.html). Section 1 presents the purpose and scope of the document, as well as a list of other ATBDs that describe some key VIIRS inputs for this EDR. Section 2 provides background information through a review of the relevant literature, with a particular focus on the current applications of, and approaches to retrieve, land cover information. This section also presents an experiment overview that lists the requirements for the surface type retrievals as set forth by the Integrated Program Office (IPO) in the Sensor Requirements Document (SRD), some of the principal characteristics of the VIIRS instrument, and a brief algorithm retrieval strategy. The theoretical and mathematical descriptions of the candidate algorithms are outlined in Section 3, which also includes some discussion on the processing and implementation strategies. A listing of the principal assumptions and limitations of this EDR appears in Section 4. Section 5 presents preliminary results produced from analyses performed during algorithm development. Section 6 lists the references cited in the text. Finally, Appendix A provides a pilot study that examines the potential for soil type retrievals from VIIRS. Although the retrieval of soil types is no longer a specified requirement for the Surface Type EDR, this study may provide future algorithm developers with the information necessary to classify broad soil types from future sensors such as VIIRS.

1.3 SUPPORTING VIIRS DOCUMENTATION

Other referenced documents describing in detail several EDR retrievals and/or output products that will serve as input data to the Surface Type EDR are as follows.

[V-1] Y2398—VIIRS Surface Albedo ATBD

[V-2] Y2400—VIIRS Vegetation Index ATBD

[V-3] Y2401—VIIRS Snow Cover/Depth ATBD

[V-4] Y2411—VIIRS Surface Reflectance ATBD

[V-5] Y2412—VIIRS Cloud Mask ATBD

[V-6] Y2468—VIIRS Operations Concept Document

[V-7] Y2469—VIIRS Context Level Software Architecture

[V-8] Y2470—VIIRS Interface Control Document (ICD)

[V-9] Y2474—VIIRS Land Module Level Software Architecture

[V-10] Y2483—VIIRS Land Module Level Detailed Design

[V-11] Y3236—VIIRS Software Integration and Test Plan

[V-12] Y3237—VIIRS Algorithm Verification and Validation Plan

[V-13] Y3257—VIIRS Computer Resources Requirements Document

[V-14] Y3261—VIIRS RDR to SDR Conversion ATBD

[V-15] Y3270—VIIRS System Verification and Validation Plan

[V-16] Y3279—VIIRS Land Module Level Interface Control Document

[V-17] Y3252—VIIRS Active Fires ATBD

[V-18] Y6635—VIIRS Algorithm Software Development Plan

[V-19] Y6661—VIIRS Algorithm Software Maturity Assessment

[V-20] Y7040—VIIRS Algorithm/Data Processing Technical Report

[V-21] Y7051—VIIRS Earth Gridding ATBD

1.4 REVISIONS

This is Version 4 of this ATBD and is dated May 2001. Version 3 was dated May 2000. Version 2 was dated June 1999, and Version 1 was dated October 1998. Version 0.2 was submitted for review by VIIRS Science team members in September 1998. Version 0.1 was provided in July 1998 for internal review by several VIIRS land team members. An ATBD summary was previously submitted in June 1998. This document has been minimally updated since Version 3. The only significant change is the renaming of the VIIRS bands. More extensive changes relating to quality flags and the interplay between this EDR and other EDRs in the VIIRS system will be incorporated into Version 5 of this ATBD.

2.0 BACKGROUND INFORMATION AND EXPERIMENT OVERVIEW

2.1 THE UTILITY OF COARSE SCALE LAND COVER INFORMATION

2.1.1 Global Modeling

The need for contemporary, accurate, and repeatable global land cover classifications to support global change research arises for a variety of reasons. Present land cover conditions are needed to generate fields of land cover-dependent biophysical parameters used in many current General Circulation Models (GCMs). These models can simulate atmospheric circulation and climatic variables such as temperature, rainfall, humidity, and wind under various global warming scenarios and at a fairly coarse spatial scale (Dickinson *et al.,* 1986; Sellers *et al.,* 1994). Many current GCMs are now coupled with Land Surface Parameterization (LSP) models. These depend on fairly recent compilations of global land cover such as those of Olson *et al.* (1983), Matthews (1983), Wilson and Henderson-Sellers (1985), and DeFries and Townshend (1994a). They also provide the means by which to include the fine-scale heterogeneity of land processes within the coarser grid of the GCMs. The LSPs come from a realization that vegetation and soils play an important role, both in space and time, in regulating the exchange of energy, gases, and water vapor between the biosphere and the atmosphere and, as such, should be included in GCM simulations (Charney, 1975; Dickinson, 1983). The LSPs also serve to produce databases or look-up tables of land cover dependent albedo, surface roughness, evapotranspiration, and respiration. These parameters control, respectively, the transfer of energy, momentum, mass, and latent and sensible heat between the biosphere and the lower layers of the atmosphere (Sellers *et al.,* 1994; Dickinson, 1995).

Land cover information is also an important input to biogeochemical, ecosystem and hydrological models which track the cycling of carbon, nutrients, energy and water between the biosphere and the atmosphere (Melillo *et al.,* 1993; Melillo, 1994; Running and Hunt, 1993; Nemani and Running, 1996). These models can simulate the response of terrestrial ecosystems to elevated CO_2 concentrations and/or climate change. By quantifying the net primary production of these ecosystems, for example, they can help to identify the principal sources and sinks of carbon, and their temporal and spatial variability, as well as providing improved estimates of the size of various global carbon pools. Vegetation type information is important to these models. Various plant and tree species have varied mechanisms for photosynthesis and carbon assimilation, which can be affected by different stresses. All these factors can significantly alter estimates obtained from the models (Bonan, 1995).

2.1.2 Land Cover Change

Another important aspect of these modeling activities is the inclusion of land cover change or land cover conversion. Current estimates of the release of carbon to the atmosphere due to land cover change have large uncertainties associated with them (Houghton and Woodwell, 1989). These uncertainties have obvious implications to the accurate estimation of net carbon exchange between the biosphere and the atmosphere, and thus to studies of the global carbon cycle. Land cover change can significantly affect not only the carbon balance of the planet but also global bio-diversity, nutrient cycles, land degradation, as well as local and regional meteorology (Nemani and Running, 1995). Several authors have suggested that in fact, changes in land cover may be as significant, and perhaps more so, as those resulting from an increase in greenhouse

gases at regional and local scales (Shukla *et al.,* 1990; Skole, 1994). There is currently large disagreement between estimates of the land cover conversion which has occurred in the past, is occurring now, and on the rates of change in land cover conversion (Skole and Tucker, 1993; Williams, 1994; Graetz, 1994). Clearly, if a series of consistent global land cover classifications can be generated for a period of several years, significant changes in global land cover over time, and their rates of change, could be quantitatively evaluated and many of the above uncertainties reduced.

2.1.3 Better Maps and Estimates of Rates of Change

Recent comparisons of widely available digital global land-cover classifications compiled from ground-based sources have noted large discrepancies, both in terms of the spatial distribution of different major land cover types and their actual areal extent over the globe (Townshend *et al.,* 1991; DeFries and Townshend, 1994a). This large disagreement in estimates may be due to the varied sources and methodologies used in compiling the different maps, but it is not uncommon. Other frequently used sources have also shown large differences in quantifying the amount of important cover types such as forests and grasslands of the world, for example (Williams, 1994; Graetz, 1994). It can be expected that updated global land cover classifications may also disagree because of differences in the land cover classifications schemes used, and the research activities they are designed to support (Hansen *et al.,* 2000). However, updated classifications derived from satellite data should certainly reduce the uncertainties currently present. Land cover information is a significant input for the global change modeling activities noted above. If we cannot accurately reconcile the current distributions of the principal land cover types, then the tasks of closing the carbon cycle or simulating global change, for example, become that much more difficult.

2.1.4 Support of Future Remote Sensing Missions

Finally, the importance of land cover information to support current and ongoing global change research is also evident in the algorithm design for a variety of products to be generated from data acquired from new satellite remote sensing platforms such as MODIS (Running *et al.,* 1994; Justice *et al*. 1998). The algorithms of Strahler *et al.* (1996a, 1999) are slated to produce improved land cover classifications, as well as land cover change products. The cloud masking algorithm of Ackerman *et al.* (1997) makes use of the land cover database of Olson *et al.* (1983), and that of Loveland *et al.* (1991) over North America, to provide cloud-free information for a great number of users downstream. The algorithm of Vermote (1996) utilizes land cover information generated using methodology developed by Running and Nemani (1995) to couple surface bidirectional effects with atmospheric scattering and absorption effects and thus improve his atmospheric correction results. The six biome classification approach of Running and Nemani (1995) is also used to produce look-up tables of parameters such as Leaf Area Index (LAI) and the fraction of Absorbed Photosynthetically Active Radiation (FAPAR), which are then used to estimate Net Primary Productivity (NPP) and net photosynthesis (Running *et al.,* 1996). Additionally, several VIIRS EDRs will make use of surface type information for their retrievals. Thus, the improved characterization of global land cover is a very important component to create new and improved data sets to study global change.

While global change studies have been performed for quite some time, a study of global change which is truly “global” in scope can only be carried out through the use of satellite remote

sensing. These satellites provide a nearly continuous and quantitative record, at spatial scales ranging from the local, to the continental and global. Several decadal global data sets acquired from satellite remote sensing platforms have been compiled with a sufficient length to allow us to begin addressing some of the principal contemporary questions of global change, and to make the goal of improving global land cover information an achievable one (e.g., Los *et al.,* 1994; Townshend, 1994; James and Kalluri, 1994). Future satellite systems, such as those planned for EOS and NPOESS, will continue to provide the much improved data needed to further enhance our knowledge of the Earth's systems. The data will reduce the uncertainties which are currently associated with many areas of global change research, including land cover.

There are many other areas where surface type information is important, and at times critical. These include urban planning, fire, disaster, and deforestation monitoring, national park delineation and assessment, environmental conservation, crop health and yield estimation, and military operations. In most of these cases, however, the spatial resolution needed for accurate assessments is much finer than that provided by meteorological satellites such as the AVHRR. Nevertheless, the data from these satellites can help to identify "hot spots" or broad areas where significant changes have occurred, or are occurring. These areas can then be targeted for further study using high resolution satellites such as TM or SPOT.

2.2 APPROACHES FOR LAND COVER CLASSIFICATIONS

2.2.1 Maps and Digital Maps Compiled from Ground-based Sources

Historically, land cover classifications have been performed from ground surveys and/or other previous land maps and the mapping or delineation of land cover types has typically been made by reference to climate, physiognomic characteristics, floristic composition, or geographical location (Mueller-Dombois, 1984; Prentice, 1990). Several important points can be made about many of these classifications. First, they are subjective in that they reflect the biases of the compilers and the variety of sources they depend on. Second, they offer only qualitative information that is not very useful for input to computerized models of global change, which require quantitative surface information. Because of the varied methods, classification schemes, and age of sources, it is not always clear whether the maps reflect the potential or actual vegetation cover. An exception is the case of bioclimatic classification. Finally, because they have typically been reproduced on paper maps, updates or changes have been difficult to implement with any regularity.

Several digital maps of global vegetation (e.g., Olson *et al.,* 1983; Matthews, 1983; Wilson and Henderson-Sellers, 1985) have been compiled from a variety of ground-based sources, including maps and atlases. As previously noted, these have typically disagreed both in terms of the actual land cover present, as well as the areal extent of particular biomes (Townshend *et al.,* 1991; DeFries and Townshend, 1994a). While the above databases have been used extensively to support climate change studies, they are somewhat influenced by the decisions and choices of the compilers, as well as the quality of their sources.

2.2.2 Global Land Cover Classifications with AVHRR Data

Satellite remote sensing provides a synoptic view of the Earth and is able to perform consistent and repetitive quantitative measurements of many terrestrial processes at a variety of spatial

scales. It has been, and is currently being, explored as an attractive alternative for *actual* global land cover classifications (Tucker, 1985; Townshend *et al.,* 1987; Loveland *et al.,* 1991; Loveland and Belward, 1997; Ehrlich and Lambin, 1996; Running *et al.,* 1995, DeFries and Townshend, 1994b; DeFries *et al.,* 1995a, 1998a; Hansen *et al.,* 2000).

The above studies used remotely sensed spectral data acquired from the Advanced Very High Resolution Radiometer (AVHRR) instrument flown onboard the NOAA-7 to 14 satellite series, coupled with their temporal evolution, to separate land cover classes at the continental and global scales. These classifications have typically been based on the variability as a function of cover type of the Normalized Difference Vegetation Index (NDVI). This index, defined as the difference of the solar energy reflected from surfaces in the near-infrared and red portions of the electromagnetic spectrum divided by their sum, is now recognized as a broad indicator of surface "greenness," photosynthetic activity, and canopy phenology (Asrar *et al.,* 1984; Justice *et al.,* 1985; Daughtry *et al.,* 1992).

The approach of Loveland *et al.* (1991), and Loveland and Belward (1997), is essentially based on utilizing 12 months of NDVI data with an unsupervised classification algorithm (Figure 1a). A large database of ancillary information is used as an aid for the human interpretation of the results. While these results are indeed impressive, and the variety of output products in all likelihood is very useful for the global change community, the accurate separation of a large number of land cover classes at the global scale from just 12 months of NDVI information is really not justified without heavy human involvement and over-reliance on the ancillary data.

Other current techniques, such as those of DeFries *et al.* (1995a, 1998a), Running *et al.* (1995), Nemani and Running (1997), and Hansen *et al.* (2000), have also used the NDVI but in conjunction with information from the individual spectral bands of AVHRR, including those in the thermal wavelength region, to improve the performance of remotely sensed global land cover classifications. Ehrlich and Lambin (1996) and Lambin and Ehrlich (1996) have suggested that using a ratio of surface temperature to NDVI may improve continental scale land cover classifications beyond the use of just NDVI. Lloyd (1990) has proposed that metrics which describe the temporal evolution of NDVI for various cover types may provide additional information to resolve the different cover types.

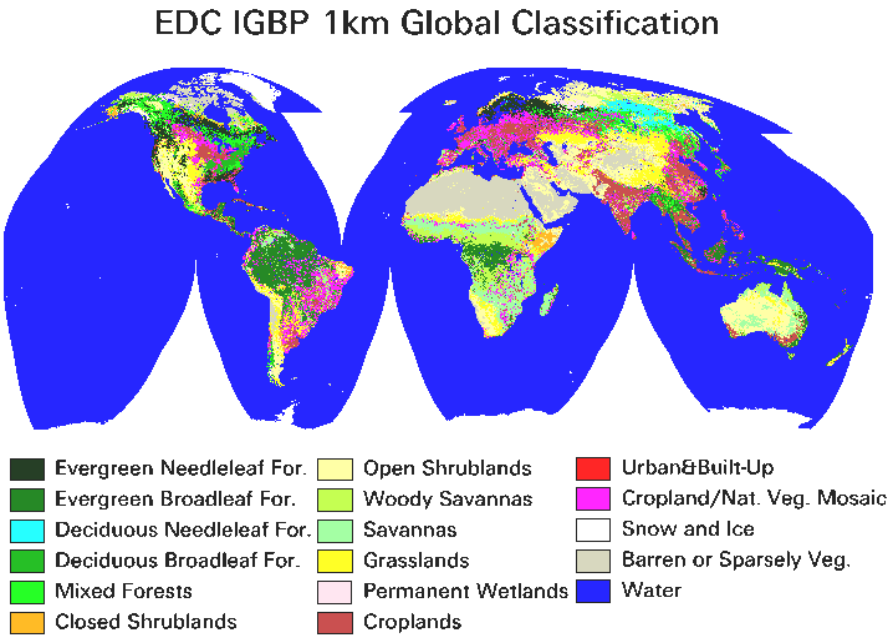
The approach developed and implemented by Running *et al.* (1995), and Nemani and Running (1997), in contrast to that of Loveland *et al.* (1991), is appealing because of its simplicity. In this approach static thresholds or decision rules for NDVI, land surface temperature, and the spectral data from AVHRR are derived at the global scale based on salient vegetation physiognomic characteristics such as leaf type and longevity. The decision rules are then used within a simple hierarchical classification scheme to yield six to eight biome types. The results generally replicate vegetation patterns at the continental to global scales and compare quite well with other available land cover data sets such as the North America classification of Loveland *et al.* (1991). However, the use of static thresholds, while robust when using data for a single year, cannot really be expected to yield consistently accurate results on all continents and for multi-year data sets, particularly within an operational scheme. The thresholds, if used in an operational setting, would have to be continuously updated and/or modified to better reflect inter-annual variability and would not be easily implemented in an automated fashion.

The work of DeFries *et al.* (1995a, 1998a) and Hansen *et al.* (2000) is a supervised classification approach which relies on a data set of carefully screened global training data and is nearly completely automated (Figure 1b). The approach is very flexible and appears to have the most promise to produce a primarily objective and operational global land cover classification. Their training data can be used at a variety of spatial resolutions, with a variety of classification algorithms, supplemented and refined with data from more land cover types. The training data can also be used for classification using remotely sensed input layers in the form of NDVI composites, spectral and temporal metrics (such as those suggested by Lloyd, 1990), and surface temperature-NDVI ratios (DeFries *et al.,* 1995a, 1998a; Hansen *et al.,* 1996, 2000).

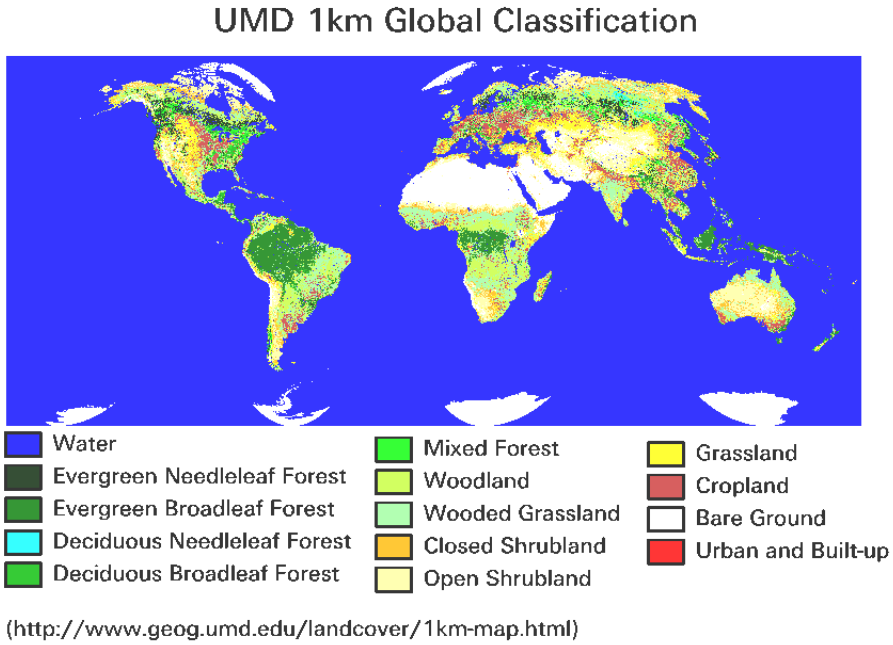
Methods to improve the accuracy of current or planned global land cover classifications attempt to capitalize on two types of pattern recognition algorithms that are fairly new to the field of remote sensing: decision trees and neural networks (DeFries *et al.,* 1998a; Hansen *et al.,* 1996, 2000; Strahler *et al*., 1996a, 1999; Friedl and Brodley, 1997). In their simplest form, decision tree classifiers successively partition the input data into more and more homogeneous subsets by producing optimal rules or decisions, also called nodes, which maximize the information gained and thus minimize the error rates in the branches of the tree. Typically, the tree overfits the data and branches or leaves with higher error rates are then pruned to produce the final output. Each final leaf is then the result of following a set of mutually exclusive decision rules down the tree.

Neural networks are another class of machine learning algorithms which are designed to intuitively resemble the human brain and can vary significantly in terms of complexity (several examples are given in Strahler *et al.* 1996a). The basic unit of the neural network is the perceptron, several of which can be networked together in one or more layers, with results from one perceptron feeding the calculations of another and vice versa. These calculations are typically based on the derivation and successive adjustment of weights as each new case is ingested and errors are detected. The final product is then produced from the set of weights that minimize the misclassification errors (Weiss and Kulikowski, 1991).

Decision trees and neural networks are particularly well suited to global land cover applications because they are non-parametric, in that they do not make any implicit assumptions about Gaussian or normal distributions in the input data, as a statistical classifier such as a Maximum Likelihood Classifier (MLC) would. Both these techniques have in fact been shown to be much superior in classification accuracies to MLCs at a variety of spatial scales (Strahler *et al*., 1996a, 1999; Friedl and Brodley, 1997), and with a variety of input data (DeFries *et al.,* 1998a; Hansen *et al.,* 1996). As has been shown in DeFries *et al.* (1998a), training data, and particularly global scale training data, tend towards non-gaussian distributions in multi-spectral space. Decision trees and neural networks are well suited to this type of problem and both algorithms tend to produce comparable classification accuracies when tested with the same remotely-sensed data inputs (Strahler *et al.* 1996a, 1999). However, decision trees are typically less computationally expensive than most neural networks (Weiss and Kulikowski 1991) and, by virtue of their hierarchical structure, also provide analysts and users with a simpler yet robust method to interpret, test, and analyze their results (Hansen *et al.,* 1996; Friedl and Brodley, 1997).



a) (http://edcwww.cr.usgs.gov/landdaac/glcc/globe_int.html)



(http://www.geog.umd.edu/landcover/1km-map.html)

b)

Figure 1. AVHRR global land cover classifications for 1992-1993. a) EDC 1km product with IGBP classes produced using an unsupervised clustering algorithm (Loveland and Belward 1997); b) University of Maryland 1km product derived from a supervised classification using a decision tree approach (Hansen *et al.* 2000). These are the only currently available global 1km land cover products and will serve as MODIS at-launch products. NOTE: Classification and color schemes are different for both a) and b).

Recently, a new paradigm for the classification of global land cover has been proposed (DeFries *et al.,* 1995b, 1999). Linear mixture models have been used to generate continuous fields of vegetation characteristics, based on the specification of fairly "pure" end members of leaf type, or amount of woody material, for example (DeFries *et al.,* 1995b, 1999). Figure 2 shows an example of such a product, developed from 1km AVHRR data (DeFries *et al*. 1999), where the value for each pixel corresponds to the percentage of tree cover within the cell. This product, along with a suite of other global continuous fields products, are slated to be produced from MODIS data (Townshend 1999). These products are found to more closely represent natural gradients in vegetation characteristics, as opposed to the classification of cover types into discrete values and in fact can potentially be more useful to global modelers than stratifications by land cover, as has previously been done. They can be, however, sensitive to the accurate description and/or purity of the end members, a problem that can be significant at coarse scales.

2.2.3 Current Limitations

The classification accuracies that have been reported in the literature range from 70 to 90 percent (DeFries *et al.,* 1995a, 1998a; Hansen *et al.,* 1996, 2000; Strahler *et al,* 1999; Friedl and Brodley, 1997). However, these are typically obtained from a sample of unseen cases from the same data used to train the classifier. The accuracies could be expected to be lower when applied to truly independent validation data. The MODIS land cover product (Strahler *et al*., 1996a, 1999) is expected to have an achievable accuracy around 80 percent. The correct typing probability for the 1 km global land cover classification of Loveland and Belward (1997) was expected to reach 85 percent but is in fact much lower than this at around 67 percent (Scepan 1999).

The limitations to achieving higher recognition accuracies have been discussed in both DeFries *et al.* (1998a) and Hansen *et al.* (2000), particularly in terms of the data quality of the input AVHRR data. These data are found to sometimes include artifacts of processing, substantial radiometric noise and/or geolocation errors. Likewise, the limited spectral coverage of the AVHRR may not provide sufficient surface information to separate similar land cover types. In most cases, the limitations are simply due to the fact that many land cover classes, particularly at the coarse spatial scale, show more natural intra-class than inter-class variability. This problem can be exacerbated when kilometer scale pixels containing several land cover classes are considered, further suggesting the need for a linear unmixing approach.

It should be noted that the NDVI is not directly related to surface or canopy structural or architectural attributes. Thus, land cover types which exhibit similar NDVI temporal signals, yet are structurally different, may still be difficult to separate. The NDVI, and the maximum value compositing approach used to produce monthly NDVI composites, are designed to reduce variations introduced by cloud cover, atmospheric and topographic effects, as well as the viewing and illumination geometries (Holben *et al.,* 1986). However, the spectral and temporal signals used for land cover classifications may still be significantly contaminated by these effects. This affects both the overall accuracy of the classifications and the ability to satisfactorily discriminate between different cover types.

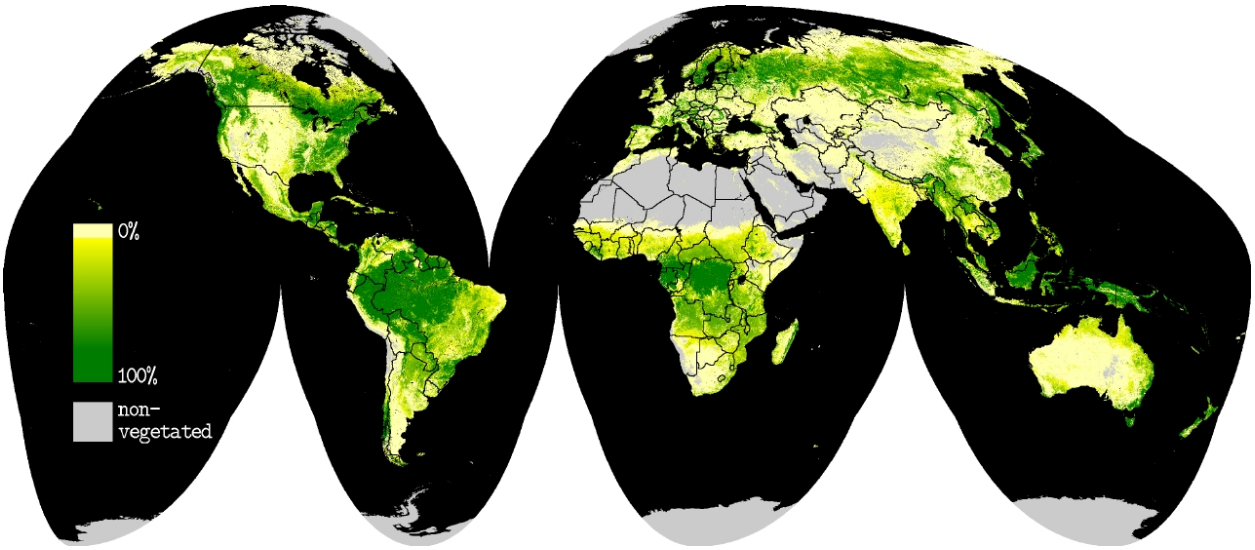


Figure 2. Prototype AVHRR 1km product for 1992-1993 showing the percentage of tree cover at the global scale (DeFries *et al.* 1998b). The percent tree cover may be underestimated in areas with significant cloud cover throughout the year. This is a recent compilation showing products that will be available in the future from polar orbiting satellites such as MODIS and VIIRS.

The use of additional metrics obtained from the individual spectral bands of AVHRR has been proposed by DeFries *et al.* (1995a, 1998). This provides greater information content and is useful in enhancing the separation of similar cover types. However, due to the large scan angle of the instrument, both the atmospheric and sun-target-sensor geometry effects may become more significant here than when using an NDVI-only approach.

Clearly, better corrections are needed for both atmospheric perturbations as well as those introduced by the non-Lambertian behavior of terrestrial surfaces. These corrections are used for MODIS and will be included in the NPOESS processing stream. The data available from EOS and NPOESS sensors such as MODIS and VIIRS will also be of significantly superior radiometric performance and stability. They will have greater spectral coverage, with much improved geolocation and registration. The substantial improvements in these data sets should provide improved land cover products beyond the accuracies that are currently achievable from AVHRR data.

To our knowledge, the land cover products generated from VIIRS data will be the first coarse scale land cover classifications that will be produced in an *operational* mode. Therefore, the approach taken needs to be robust and computationally efficient, yet fairly simplified from what is proposed for MODIS. For example, several studies have suggested that the bidirectional domain of remote sensing may contain surface information that could be used in improving land cover classifications (Abuelgasim *et al.*, 1996; Wu *et al*., 1995; Walthall and Brown de Colstoun, 1997). In fact, the MODIS land cover product may use BRDF products produced from combined MODIS and MISR data (Strahler *et al*., 1996b). Because of the computational load created by this approach, the need for two satellite sensors to obtain a sufficient sample of the surface BRDF, and the unproven benefits of BRDF information for classifications, this type of approach will not presently be implemented for the VIIRS Surface Type EDR.

2.3 EXPERIMENT OVERVIEW

2.3.1 Surface Type EDR Requirements

The SRD requirements for the Surface Type Environmental Data Record (EDR) demand that the VIIRS data and algorithm be able to classify the terrestrial surface for both moderate and fine horizontal cell sizes (HCS). The classification will be performed under clear conditions only, and within an operational computing environment. The land surface will be classified according to the 17 surface types specified by the International Geosphere Biosphere Programme (IGBP). The overall typing accuracy for these 17 land cover types should be at least 70 percent and the spatial resolution 20 km or better for the moderate HCS product and 1 km or better for the fine HCS product (Table 1). We expect that a typing accuracy of 90 percent will be achievable in the NPOESS era.

Additionally, the SRD states that "Estimation of the percentage of vegetation cover per type in each cell is an objective." The accuracy and precision of this vegetation cover objective requirement is 2 % and 0.1 %, respectively. This requirement implies being able to report the percentage cover of each IGBP class present within the cell to the specified accuracy and precision, and NOT the percentage of vegetation cover present in the cell.

2.3.2 Instrument Characteristics

The principal spectral, spatial, radiometric, and temporal characteristics of the VIIRS instrument are described in the VIIRS Sensor Specification. Of particular relevance to this EDR is the spectral coverage (Table 2), with VIIRS bands M4 (545-565 nm), I1 (620-670 nm), I2 (846-885) nm), I3 (1580-1640 nm) and M11 (2105-2155 nm) being the most critical in the visible/near and middle infrared portions of the spectrum. The Brightness Temperatures (BT) measured in VIIRS thermal channels I4 (3.55-3.93 μm), M15 (10.26-11.26 μm) and M16 (11.54-12.49 μm) will also be quite important. The temporal resolution or global repeat cycle of the NPOESS platforms, which are yet to be determined, will also be another important instrument parameter, particularly since the classification will be built from time composited data such as is currently done with AVHRR.

Land cover classifications have been successfully performed for quite some time using the five broad spectral bands of the AVHRR as well as the seven spectral bands of the LANDSAT Thematic Mapper (see Table 2). The MODIS land cover algorithm (Strahler *et al*., 1996a) is slated to use the seven "land" bands of MODIS to produce global land cover classifications to an expected 80 percent accuracy. VIIRS bands M4, I1, I2, I3, and M11 approximately correspond to MODIS bands 4, 1, 2, 6, and 7 and their center wavelengths are well within bands 2, 3, 4, 5, and 7 of the LANDSAT TM.

Table 1. SRD Requirements for the Surface Type EDR

Parameter No.	Parameter Name	Thresholds	Objectives
	a) Horizontal Cell Size		
V40.6.4-1	1) Moderate, worst case	20 km	1 km
V40.6.4-2	2) Fine, at nadir	1 km	0.25 km
V40.6.4-3	b) Horizontal Reporting Interval	(TBD)	(TBD)
V40.6.4-13	c) Horizontal Coverage	Land	Land
	d) Measurement Range		
V40.6.4-6	1) Surface Type	17 IGBP classes	17 IGBP classes
V40.6.4-7	2) Vegetation Cover	N/A	0 - 100%
V40.6.4-8	e) Measurement Accuracy (veg. cover)	N/A	2%
V40.6.4-9	f) Measurement Precision (veg. cover)	N/A	0.1%
	g) Correct Typing Probability (surface type)	70%	(TBD)
V40.6.4-14	1) Moderate, worst case	70 % at (TBS) % confidence level	(TBD) at (TBS) % confidence level
V40.6.4-15	2) Fine, at nadir	70 %	(TBD)
V40.6.4-11	h) Mapping Uncertainty	5 km	1 km
	i) Maximum Local Average Revisit Time	24 hrs	3 hrs
	j) Maximum Local Refresh	(TBD)	(TBD)
V40.6.4-12	k) Minimum Swath Width (All other EDR thresholds met)	3000 km (TBR)	(TBR)

Sensor Requirements Document, Section 3.2.1.1.1.16
(TBD: To Be Decided, TBR: To Be Resolved)

Figures 3, 4, and 5 show the spectra for several typical cover types in different regions of the electromagnetic spectrum. The figures also show the relative position of several of the nominal VIIRS spectral bands. As can be seen, the current spectral coverage for vegetation, soils and snow is quite good, capturing areas of optimal separability for these cover types.

It should be stressed that the spectra for the vegetated cover types presented here are measured in the laboratory, on a single date, for piles of leaves or needles, and that natural cover types may show increased separability in these spectral regions because of canopy and other structural differences. Additionally, the change of these spectra with time (e.g., during a growing season) would tend to increase the differences at different times of the year. If the temporal evolution of the target were represented with temporal metrics such as those developed by DeFries *et al.* (1995, 1998a), these differences would be more apparent than seen in these figures.

Table 2. Intercomparison of the Spectral Bands of VIIRS, MODIS, AVHRR, and TM

VIIRS			MODIS Equivalent			AVHRR Equivalent (AVHRR-3)			TM Equivalent (Landsat 4 & 5)		
VIIRS Band	Spectral Range (um)	Nadir HSR (m)	Band	Range	HSR	Band	Range	HSR	Band	Range	HSR
M1	0.402-0.422	750	8	0.405-0.420	1000						
M2	0.435-0.455	750	9	0.438-0.448	1000						
M3	0.478-0.498	750	10	0.483-0.493	1000				1	0.450-0.520	30
M4	0.545-0.565	750	4	0.545-0.565	500				2	0.520-0.600	30
I1	0.620-0.670	375	1	SAME	250	1	0.572-0.703	1100	3	0.630-0.690	30
M5	0.662-0.682	750	1	0.620-0.670	250	1	0.572-0.703	1100	3	0.630-0.690	30
M6	0.739-0.754	750	15	0.743-0.753	1000						
M7/I2	0.846-0.885	750/375	2	0.841-0.876	250	2	0.720-1.000	1100	4	0.760-0.900	30
M8	1.230-1.250	750	5	SAME	500						
M9	1.371-1.386	750	26	1.360-1.390	1000						
M10/I3	1.580-1.640	750/375	6	1.628-1.652	500	3a	SAME	1100	5	1.550-1.750	30
M11	2.225-2.275	750	7	2.105-2.155	500				7	2.080-2.350	30
I4	3.550-3.930	375	20	3.660-3.840	1000	3b	SAME	1100			
M12	3.660-3.840	750	20	SAME	1000	3b	3.550-3.930	1100			
M13	3.973-4.128	750	21-23	3.929-4.080	1000						
M14	8.400-8.700	750	29	SAME	1000						
M15	10.26-11.26	750	31	10.78-11.28	1000	4	10.3-11.3	1100	6	10.40-12.50	120
I5	10.50-12.40	375	32	11.77-12.27	1000	5	11.5-12.5	1100	7	10.40-12.50	120
M16	11.54-12.49	750	32	11.77-12.27	1000	5	11.5-12.5	1100	8	10.40-12.50	120

Although MODIS band 3 (459-479 nm) is not currently included in the nominal VIIRS band set, usage of VIIRS bands M2 (433-453 nm) and/or M3 (480-500 nm) instead could provide a substitute to this band because the land surface information content of these bands with regards to land cover is highly correlated to that of MODIS 3 (Figure 3). Also, VIIRS band M9 would seem to offer little additional information for land cover classification, except in the case of snow (Figure 4). It is suggested that the inclusion of MODIS band M8 (1230-1250 nm) (Figure 4) may provide additional information for both land cover characterization and albedo estimations, but it is not critical. The true "critical" bands for this EDR are VIIRS bands M4, I1, I2, I3, and M11, and obviously those needed for EDRs whose outputs are inputs to the surface type EDR.

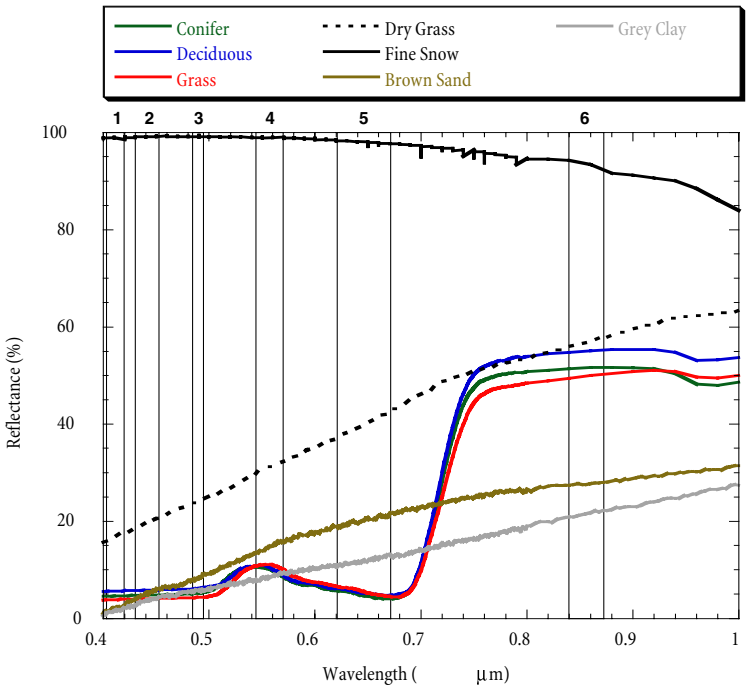


Figure 3. Laboratory measured spectra for several land cover types from 400 to 1000 nm. VIIRS nominal bands M1 through I2 are approximately shown as numbered boxes. (Spectra are available through the ASTER spectral library http://speclib.jpl.nasa.gov).

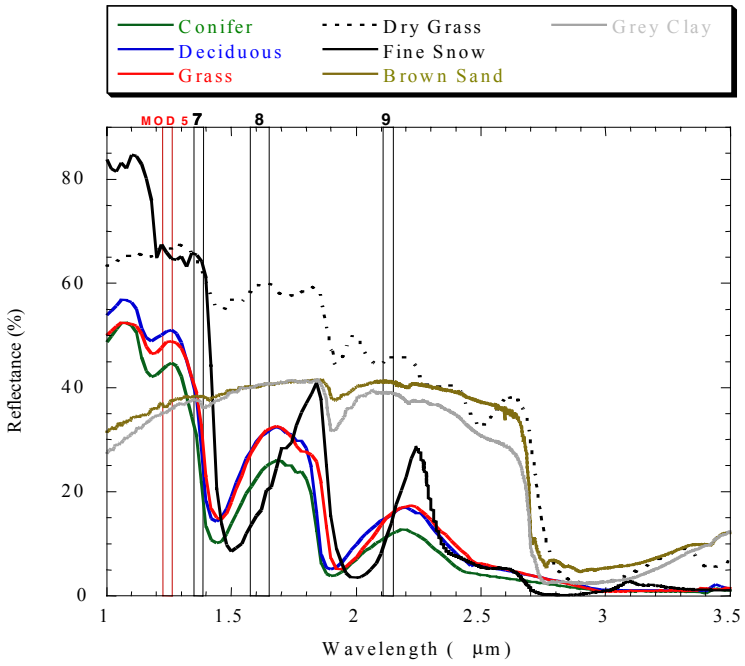


Figure 4. Spectra for same cover types as Figure 3, but for wavelengths from 1 to 3.5 µm. VIIRS bands M9, I3, and M11 are shown as black boxes. MODIS band 5 is shown as the box on the left (MOD5).

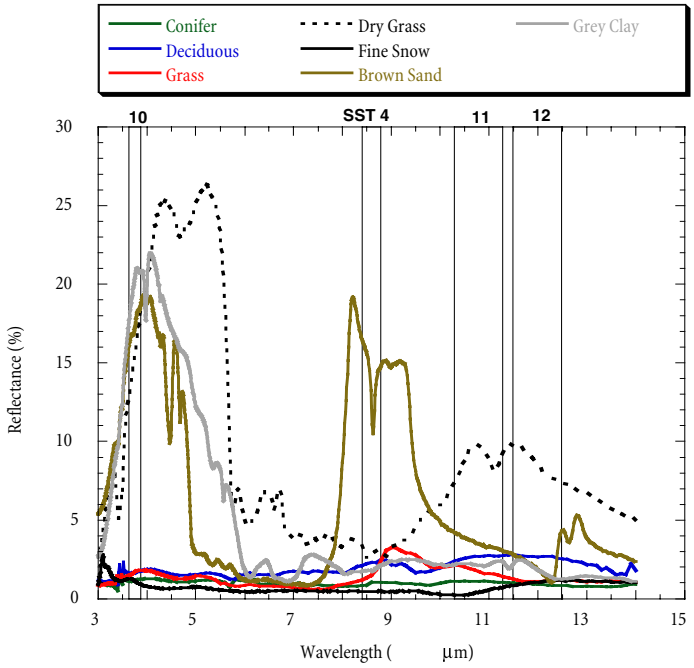


Figure 5. Same as Figures 3 and 4, but for the thermal wavelength region. VIIRS bands I4, M15, and M16 are shown as well as the 8.55 µm band (M14).

Studies have shown a potential for improving the discrimination of forest from non-forest types such as savanna by using AVHRR band 3 (Laporte *et al.,* 1995; 1998). Although this band (I4 (3.55-3.93 µm)) (Figure 5) is located in a spectral region containing both reflected solar and emitted terrestrial energy, these results suggest that this band should also be included in the classification.

In the thermal region, a prominent feature from 8 to 10 microns can be easily seen for the brown sand spectrum (Figure 5). This spectral feature is related to the amount of quartz or sand within the soil (Salisbury and D'Aria, 1992a,b), and is not present in the spectrum of the gray clay that contains relatively little sand. It is hoped that the information contained within this band may thus be helpful in separating sandy soils, clay, and loam directly from the VIIRS data (see Appendix A). These are issues that are still under investigation and will need to be fully resolved before they can be included within an operational Surface Type algorithm.

As can be seen, the spectral coverage of VIIRS with regards to this EDR is very good, and offers a good compromise between a hyperspectral type approach with many highly correlated bands, and the AVHRR. It also offers significant flexibility in the number of bands that can be used in the algorithm because individual spectral bands that are not critical to the successful performance of this EDR can still be included to provide additional land cover information.

2.3.3 Retrieval Strategy

The approach that will be taken to achieve threshold requirements, and to approach objective requirements, will be essentially a supervised classification approach, in that representative

training samples from the land cover classes will be used to train our decision tree classification algorithm. These training data have been developed by scientists at the University of Maryland, and used extensively to classify AVHRR data at spatial resolutions ranging from 1 km to 1 degree (DeFries *et al.,* 1995a, 1998a; Hansen *et al.,* 1996, 2000). Other sets of training data were produced in-house from the 1 km AVHRR global land cover classification performed at the EROS Data Center (EDC) (Loveland and Belward, 1997).

The operational Surface Type EDR will be developed around a global 1 km VIIRS Quarterly Surface Types Intermediate Product (IP) which will be produced every three months, and which will meet the requirements for both the moderate and fine HCS products. For this product, temporal metrics describing the temporal evolution of training data over the 12 previous months will be used for the classification, following the methodology developed at the continental scale by Tucker *et al*. (1985) and Townshend *et al.* (1987), and then extended to the global scale by DeFries *et al.* (1995a; 1998a) and Hansen *et al.* (2000). During each successive three month period, the VIIRS Quarterly Surface Types IP will be re-delivered for every VIIRS orbit in conjunction with the current VIIRS Snow Cover/Depth and Active Fires EDRs. The fraction of green vegetation cover present per cell will also be provided as a part of this EDR, following Gutman and Ignatov (1998). This suite of products will provide users with current surface condition information associated with each cell and surface type of the Quarterly Surface Type product. This will eliminate the need to perform a global classification operationally, which would create much redundant information, as kilometer scale changes from one surface type to the next do not generally occur on a daily basis on a large area of the Earth, except for exceptional or catastrophic events such as very large fires, floods, or volcanic eruptions.

A Quarterly Continuous Fields IP tailored after the MODIS continuous fields products (Townshend 1999) will also be produced every three months using the same metrics and training data used in the generation of the Quarterly Surface Types IP. The VIIRS Surface Types-Biomes and Olson IPs will also be produced every three months by aggregating the IGBP classes into the appropriate surface types. These products are to be used internally by several VIIRS EDRs.

The Surface Type EDR is involved in the production and/or use of four additional, Earth-gridded, intermediate products. It requires as input the Monthly Vegetation Index IP (MVI), the Monthly Surface Reflectance IP (MSR), and the Monthly Brightness Temperature ATBD (MBT). These three products are required to produce the input metrics for the decision tree, and represent, respectively, the maximum value composite of cloud-free NDVI over a month, the average cloud-free surface reflectance over a month, and the average cloud-free TOA brightness temperature over a month. Three months of each are accumulated for use in generating the Quarterly Surface Types IP. The generation of all three products is described in *[V-21]*.

The Surface Type EDR must in turn be used to generate a Forest Mask IP, for use by the Snow Cover/Depth EDR. This mask is produced in the Build SDR module, along with the other Earth-gridded products. It is generated using the most recent output of the Surface Type EDR, which includes clear delineation of forest from other types of land cover.

3.0 ALGORITHM DESCRIPTION

3.1 SURFACE TYPES

One of the greatest difficulties in current global land cover classification research is the lack of consensus regarding the classification scheme. Broad scale global modeling can be satisfied with the characterization of six to eight simple biome types such as forests, grasslands, and crops, but regional scale biogeography demands more information of species assemblages, forest/grassland mixtures, for example (Running *et al.,* 1995). Difficulties also exist in reconciling the output data needs with the limitations of remotely sensed data, in other words, designing a classification scheme whose land cover classes, and number of classes, are potentially retrievable from coarse scale spectral and/or temporal data. The EOS equivalent of the VIIRS Surface Type EDR, MODIS Product MOD-12, is currently slated to use the 17-type classification scheme proposed by the International Geosphere-Biosphere Programme (IGBP). The scheme is currently used in several well-known global land cover classifications, notably Loveland and Belward (1997).

Table 3 shows the definitions for the 17 IGBP classes required for this EDR as available from the MODIS Land Cover Product ATBD (Strahler *et al*., 1996a). These classes encompass 11 classes of natural vegetation (classes 1 to 11 in Table 3), 3 classes of developed and mosaic lands (Classes 12, 13 and 14), and 3 classes of non-vegetated land (Classes 15, 16, and 17). This classification scheme is generally perceived to be achievable for both current and future remote sensing systems. However, as discussed by Loveland *et al*. (1999), the IGBP classification scheme still has some shortcomings. For example, a class for tundra ecosystems is not included, implying that these surface types will be included within one or more of the IGBP classes. Second, under the current definition of the Forest classes and the Wetlands classes, large portions of the Wetlands class could be contained within the different forest classes, considering that many Swamps have a canopy cover >60%. In addition, current classifications (Hansen *et al.*, 2000; Loveland and Belward, 1997) use static vector data from the Digital Chart of the World (DCW) to delineate urban areas, principally because these areas cannot effectively be classified from daytime, remotely sensed data. Cities comprise a mosaic of man-made materials, water, and perennial and annual vegetation, the distribution of which varies both within, and across, continents. Whether improved satellite data will allow the direct retrieval of urban areas at a global scale is not yet known. Finally, the IGBP Natural Vegetation/Croplands Mosaics class may not be useful for coarse scale modeling activities because these models typically require aggregated surface type information at coarser scales than 1 km. It may in fact cause problems when this aggregation is performed if this class is the dominant class within a coarser cell.

The previous version of this ATBD had presented a modified classification scheme based on the work of DeFries *et al.* (1998) and Hansen *et al.* (2000), the classes required by the IPO in the previous version of the SRD, and the IGBP classes. This scheme was also presented during our land TIMs. In any case, we feel that the IGBP classification scheme is mostly correct and encompasses surface types that should be retrievable from future operational remote sensing platforms, as it has successfully been implemented using AVHRR data. Better future data with more spectral bands may allow more classes than the current 17 to be retrieved. However, we caution the reader that, particularly in light of the limitations of the AVHRR data listed in section 2.2.3, classification schemes that present the retrieval of many more classes than the IGBP scheme, or who claim to be able to separate many sub-classes from each IGBP class from AVHRR data (e.g. Muchoney *et al.* 2000), should be viewed with a certain degree of skepticism.

Table 3. IGBP surface type definitions (from Strahler *et al.*, 1996a).

IGBP Surface Type Names	Definition
1) Evergreen Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
2) Evergreen Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees and shrubs remain green year round. Canopy is never without green foliage.
3) Deciduous Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of seasonal needleleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
4) Deciduous Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
5) Mixed Forests	Lands dominated by trees with a percent cover >60% and height exceeding 2 meters. Consists of tree communities with interspersed mixtures or mosaics of the other four forest types. None of the forest types exceeds 60% of landscape.
6) Closed Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover >60%. The shrub foliage can be either evergreen or deciduous.
7) Open Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover between 10-60%. The shrub foliage can be either evergreen or deciduous.
8) Woody Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 30-60%. The forest cover height exceeds 2 meters.
9) Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 10-30%. The forest cover height exceeds 2 meters..
10) Grasslands	Lands with herbaceous types of cover. Tree and shrub cover is less than 10%.
11) Permanent Wetlands	Lands with a permanent mixture of water and herbaceous or woody vegetation. The vegetation can be present in either salt, brackish, or fresh water.
12) Croplands	Lands covered with temporary crops followed by harvest and a bare soil period (*e.g.*, single and multiple cropping systems). Note that perennial woody crops will be classified as the appropriate forest or shrub land cover type.
13) Urban and Built-Up Lands	Lands covered by buildings and other man-made structures.
14) Cropland/Natural Vegetation Mosaics	Lands with a mosaic of croplands, forests, shrubland, and grasslands in which no one component comprises more than 60% of the landscape.
15) Snow and Ice	Lands under snow/ice cover throughout the year.
16) Barren	Lands with exposed soil, sand, rocks, or snow and never has more than 10% vegetated cover during any time of the year.
17) Water Bodies	Oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies.

Two additional surface type products will be made available for the processing of different VIIRS EDRs. The VIIRS Surface Type-Olson IP will contain the following five classes: Water, Desert, Vegetated Land, Snow/Ice, and Mountains. The VIIRS Surface Types-Biomes IP will contain six classes after Running and Nemani (1996). These will be: Grasses and Cereal Crops, Shrubs, Savannas, Broadleaf Forests, Needleleaf Forests, and Broadleaf Crops. This set will

likely be extended in Version 5 of this document to include bare soil, water, and snow, due to the global requirements of the Surface Reflectance IP *[V-4]*. The above classes will be obtained by aggregating the appropriate IGBP classes. The Mountains class may be produced through the use of a Digital Elevation Model (DEM).

3.2 PROCESSING OUTLINE

The processing chain for both the VIIRS Surface Type EDR and the VIIRS Quarterly Surface Types IP are illustrated in Figures 6 and 7, respectively.

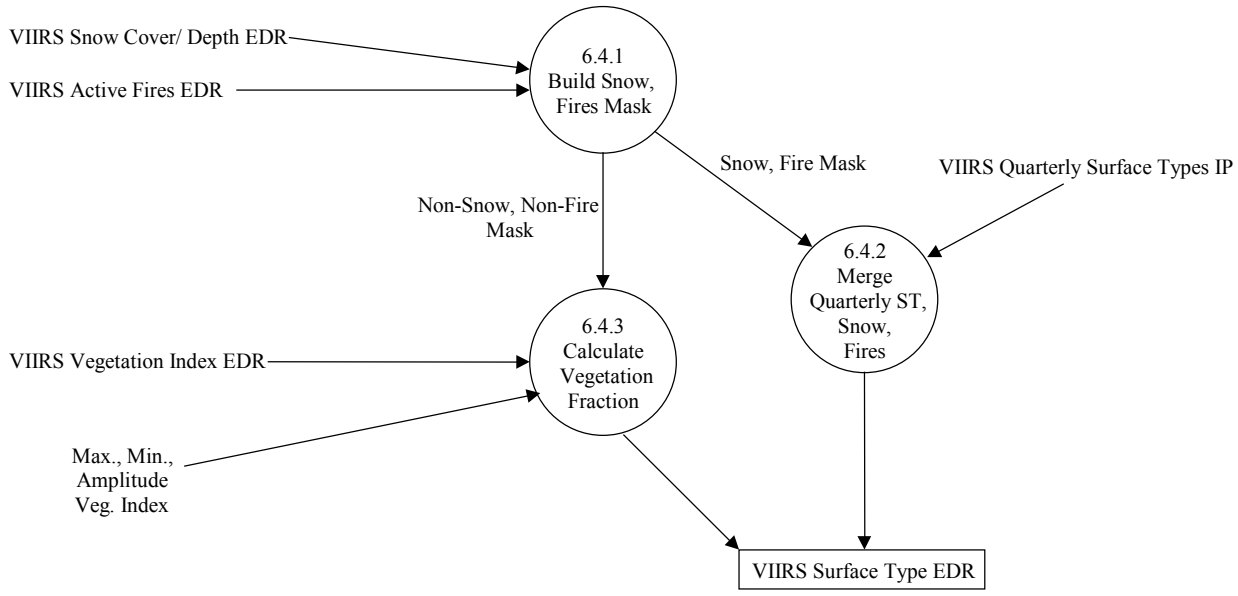


Figure 6. Flow diagram showing the processing chain for the VIIRS Surface Type EDR.

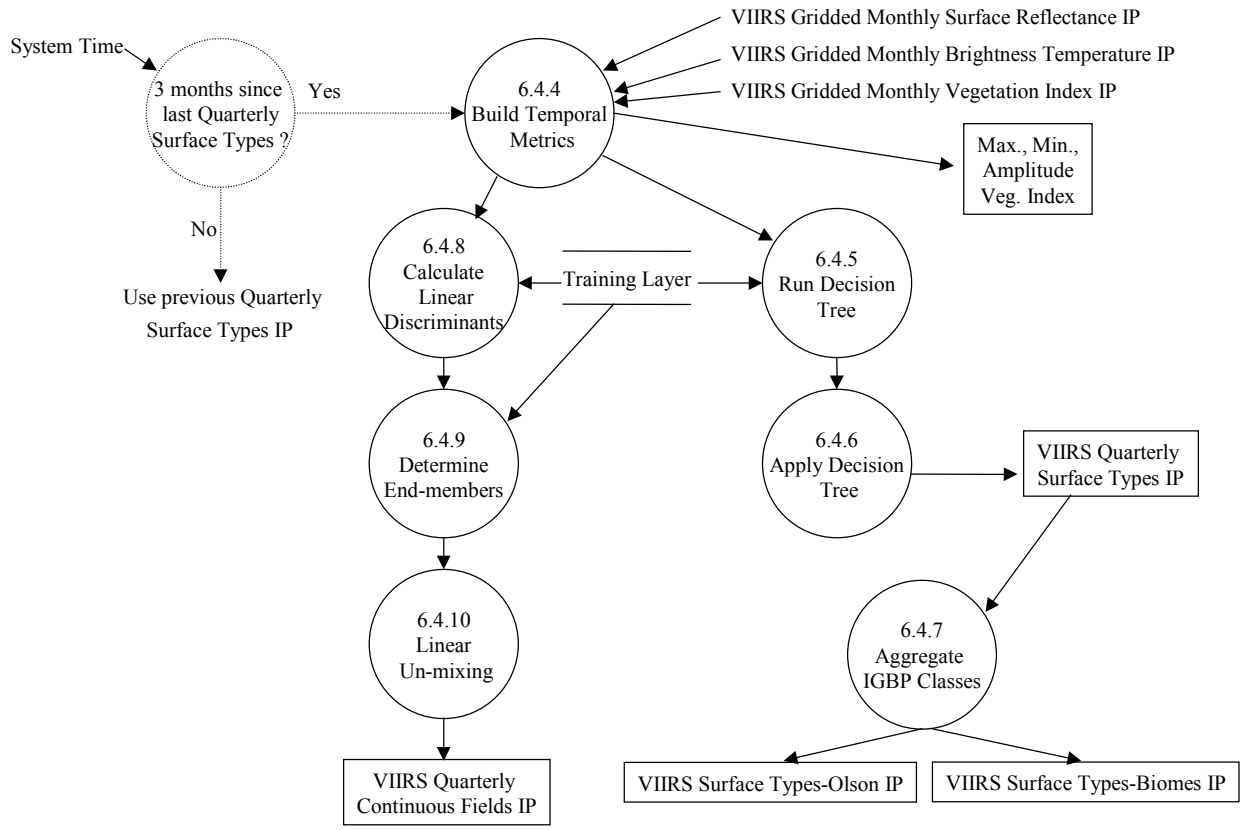


Figure 7. Flow diagram showing the processing flow for the VIIRS Quarterly Surface Types IP.

3.3 ALGORITHM INPUTS

3.3.1 VIIRS Data

The Surface Type EDR (Figure 6) will depend in large part on the output of the Quarterly Surface Types IP, as it will continuously re-deliver the most current Quarterly Surface Types product for a three month period, and for each VIIRS orbit. However, the EDR will also integrate the current VIIRS Vegetation Index *[V-2]*, Snow Cover/Depth *[V-3]*, and Active Fires *[V-17]* EDRs. A sub-module of the algorithm will calculate the green vegetation fraction in any given cell from the current VIIRS Vegetation Index (VI) EDR coupled with the Maximum, Minimum, and Amplitude (i.e. Max.–Min.) VI for the same cell as output from the Quarterly Surface Types IP from the 12 previous months of VIIRS VI data. Several indices are being generated through the VI EDR and the usage of indices such as the Soil Adjusted Vegetation Index (SAVI) (Huete *et al.,* 1988) will also be explored beyond the currently planned NDVI and EVI. The reader is directed to *[V-2]* for more current and complete details.

The inputs for the VIIRS Quarterly Surface Types IP (Figure 7) will be Earth-gridded, cloud-masked, atmospherically corrected, and monthly-composited surface reflectances for VIIRS bands M4, I1, I2, I3, and M11 (and possibly I4). These will be provided as IPs from the Earth

Gridding module through the Atmospheric Correction over Land processing chain. The gridded monthly top-of-atmosphere (TOA) brightness temperatures from VIIRS thermal bands I4, M14, M15, and M16 will also be used for the generation of the product because they provide increased discriminatory power between several land cover types (Figure 5). The earth gridded monthly vegetation index IP will also be ingested, using the NDVI as the primary vegetation index.

Temporal metrics describing the maximum, minimum, mean and amplitude for the input data layers (surface reflectances, TOA BTs, NDVI) will be produced from 12 maximum value monthly NDVI composites and serve as input to the classifiers, as suggested by DeFries *et al.* (1998a). The use of additional metrics such as those used by Hansen *et al.* (2000) is being explored at this time for future implementation. The same metrics will be used to produce both the Quarterly Surface Types IP as well as the Quarterly Continuous Fields IP.

3.3.2 Auxiliary Data

Auxiliary data are data sets that are produced outside of the NPOESS (i.e. VIIRS and Non-VIIRS instruments) processing streams. Several auxiliary data sets will be needed for the successful operation of this EDR. Products that will be needed at the time of launch will be the MODIS land/sea mask and an at-launch land cover data set. These will be obtained from a combination of the University of Maryland product (DeFries *et al.,* 1995; 1998; Hansen *et al.* 2000), an updated EDC classification (Loveland and Belward, 1997), and/or any available MODIS classifications.

The training data will be obtained from updated 1 km training areas from the work of DeFries *et al.* (1995a, 1998a) and Hansen *et al.* (2000). These training data will be supplemented by training data produced in-house from the EDC 1 km classification or other sources. These training data will be representative of the 17 IGBP land cover types (Table 3).

Because soil types are no longer required, this EDR will no longer require the use of the digital soil map of the world produced by the Food and Agriculture Organization (FAO) in Rome (FAO, 1995). This map has been compiled from exhaustive field surveys and maps over the past 30 years and is the only currently available source of digital soil information at a coarse spatial scale. The resolution of this map is currently approximately 4 km. Other users may still find this map useful.

3.4 THEORETICAL DESCRIPTION OF THE ALGORITHM

3.4.1 General Description of the Algorithm

The general approach taken for the Surface Type EDR is to use the best possible *global* land cover classification at all times. This high accuracy can only be achieved by using a temporal accumulation of VIIRS products. By coupling our Quarterly Surface Types IP with the current VIIRS Vegetation Index, Snow Cover/Depth, and Active Fires, and by providing the current green vegetation fraction, we will provide a product with the highest possible quality, but which is also updated in real-time with actual data.

Land Cover classifications are based on the assumption that different cover types will exhibit different patterns of reflected or emitted energy as a function of wavelength and/or time, or signatures, and this forms the basis for their automated recognition (Richards, 1983). The VIIRS

Quarterly Surface Types IP will be produced in a so-called "supervised" classification mode. In a supervised land cover classification, regions with particular cover types are known *a priori* and are sampled so as to be representative of the same cover types throughout an image or region. The spectral, spatial, and temporal behavior of the remotely sensed data within these training areas can then be exploited to create statistical patterns within a variety of classifiers. These patterns are used to label the remaining samples into the appropriate surface type category. By utilizing carefully selected training data from DeFries *et al* (1995a; 1998a) and Hansen *et al.* (2000), augmented by our own at the global scale, our approach will be supervised. At the global scale, we will exploit the spectral and temporal information to separate the 17 IGBP surface types.

The Quarterly Surface Type product will be based on gridded, time-composited products at 1 km resolution and will be produced and updated every three months. The accumulation of monthly composites will allow the exploitation of the temporal domain of remote sensing to enhance the separability of classes that would otherwise have similar spectral and/or NDVI signals. To facilitate data handling and processing of this large data set, the 12 consecutive monthly composites for the NDVI, and the Surface Reflectance and BTs in the VIIRS bands will be combined to produce "metrics" layers representing maximum, minimum, mean, and amplitudes for each land pixel and for each input parameter. The resulting metrics will then be input into the classifier.

The nominal classifier to be used in the production of the Quarterly Surface Types IP will be the C5.0 Decision Tree (DT) algorithm produced by Quinlan (1993). The boosting option will be used along with standard pruning parameters available within the C5.0 software. Cross-validation will be performed using several random samplings of the training data in order to provide optimal training data from which to perform the classification (Friedl and Brodley, 1997). The training data will be split into 80% training and 20% testing samples (after Strahler *et al.* 1999) and used to produce a boosted decision tree. This tree will be subsequently applied to all the remaining VIIRS metrics to produce the IP.

Alternatively, a hybrid classification scheme can be implemented. This would employ a combination of a DT classifier and a Learning Vector Quantization (LVQ) neural network classifier to potentially optimize the land cover class separation. Training pixels will be input into each of these algorithms successively, and the pixels classified correctly by each will be retained. This combined training data set will then be used within the decision tree classifier. This type of approach is still being developed and would be preferred should the results provide an improvement in typing accuracy.

We do not believe at this time that the direct retrieval of the percent cover type per cell at full resolution (objective requirement) will be achievable at the required accuracy, even with VIIRS data. In order to approach these objective requirements, however, we will implement linear mixture models within the VIIRS processing stream for the alternate production of a suite of global continuous fields products similar to that shown in Figure 2 and described in Townshend *et al.* (1999). The processing flow for these products is shown in Figure 7.

3.4.2 Mathematical Description of the Algorithm

3.4.2.1 Optimized-Learning-Rate Learning Vector Quantization (LVQ) Algorithm

In the Learning Vector Quantization (LVQ) approach, a series of so-called "codebook" vectors are assigned to each land cover class S_j. Then the codebook vector m_i which minimizes the Euclidean distance to any input sample x, is found iteratively (Kohonen, 1997). In the Optimized-Learning-Rate LVQ (OLVQ) used here, the basic LVQ algorithm is modified so that a learning rate factor $\alpha_i(t)$ is assigned to each m_i vector. The learning process follows a reward-punishment scheme and is summarized by the following set of equations:

If x is classified correctly (i.e., reinforce/reward):

$$m_c(t+1) = m_c(t) + \alpha_c(t)[x(t) - m_c(t)] \tag{1a}$$

If x is classified incorrectly (i.e., extinguish/punish):

$$m_c(t+1) = m_c(t) - \alpha_c(t)[x(t) - m_c(t)] \tag{1b}$$

and:

$$m_i(t+1) = m_i(t) \text{ for } i \neq c \tag{1c}$$

where:

$$c = \arg\min_i \{\|x - m_i\|\} \tag{2}$$

c represents the index of the nearest m_i to x, and the $0 < \alpha_i(t) < 1$ and are made to decrease with time t. Equations 1a, 1b, and 1c can be re-expressed in the form:

$$m_c(t+1) = [1 - s(t)\alpha_c(t)]\, m_c(t) + s(t)\alpha_c(t)x(t) \tag{3}$$

where s(t) = 1 if the classification is correct and s(t) = -1 if it is not.

According to Kohonen (1997), Equation 3 contains "trace" elements of x(t) in the last term, as well as of x(t'), t'=1, 2,…, t-1, from previous iterations, through $m_c(t)$. At each learning step, the last trace x(t) is scaled down by a factor $\alpha_c(t)$ and, similarly, each trace x(t-1) is scaled down by a factor $[1-s(t)\alpha_c(t)]\alpha_c(t-1)$. He further stipulates that both these scalings must be equal, i.e.:

$$\alpha_c(t) = [1 - s(t)\alpha_c(t)]\alpha_c(t-1) \tag{4}$$

If this condition holds true for all t, the $\alpha_i(t)$ will then be determined optimally by the following recursion:

$$\alpha_c(t) = \frac{\alpha_c(t-1)}{1 + s(t)\alpha_c(t-1)} \tag{5}$$

3.4.2.2 Decision Tree Algorithm

A decision tree partitions any training data set T with land cover classes C_j, for example, into more and more homogeneous subsets called nodes and leaves. In earlier decision tree algorithms, this partitioning was performed by maximizing the *gain criterion*, a measure of the information gained by sub-dividing T with a test X, and defined in Quinlan (1993) as:

$$\text{gain}(X) = \text{info}(T) - \text{info}_X(T) \tag{6}$$

where:

$$\text{info}_X(T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} \bullet \text{info}(T_i) \tag{7}$$

is a weighted sum over n subsets produced from the partitioning of T according to test X, and |T| is the total number of cases in T. Here:

$$\text{info}(T) = -\sum_{j=1}^{k} p(C_j \mid T) \bullet \log_2(p(C_j \mid T)) \tag{8}$$

where info(T) essentially represents the average amount of information needed to correctly identify a land cover class in T, and $p(C_j \mid T)$ is the probability that a particular case from a set of cases T belongs to class C_j.

In contrast, the C5.0 decision tree algorithm uses a test that maximizes the *gain ratio criterion*, which "expresses the proportion of information generated by the split that is useful, i.e., that appears useful for classification" (Quinlan, 1993, p. 23):

$$\text{gain ratio}(X) = \text{gain}(X)/\text{split info}(X) \tag{9}$$

where:

$$\text{split info}(X) = -\sum_{i=1}^{n} \frac{|T_i|}{|T|} \bullet \log_2\left(\frac{|T_i|}{|T|}\right) \tag{10}$$

This last equation is similar to Equation 8, except that it represents the additional information that is potentially gained from sub-dividing the training set into n subsets. The gain ratio criterion is used to rectify a bias of the gain criterion in favor of tests with numerous outcomes (Quinlan, 1993).

3.4.2.3 Linear Mixture Model

The linear mixture model to be used in the production of the VIIRS Quarterly Continuous Fields IP is described in full detail in Section 4 of Townshend (1999). It is based on the assumption that the reflectance of a cell is the sum of the reflectances of each component making up the cell, weighted by its respective areal coverage within the cell. This is described mathematically as:

$$R_i = \sum_{j=1}^{N} r_{ij} x_j + e_i \quad (11)$$

where R_i is the reflectance of the cell in spectral band I, r_{ij} is the reflectance of cell component j in spectral band i, x_j is the proportion of the cell that is covered by cover j, e_i is the error term, and N is the total number of components within the cell. The model further specifies that the sum of all the x_j terms should add up to unity.

3.4.2.4 Green Vegetation Fraction

The green vegetation fraction per cell that will be continuously provided with the Surface Type EDR is based on the work of Gutman and Ignatov (1998), and developed for use with the NOAA GVI product. The model that was chosen for their work is described as:

$$f_g = (NDVI - NDVI_o)/(NDVI_\infty - NDVI_o) \quad (12)$$

where f_g is the fractional green vegetation cover within a specific cell, $NDVI_o$ is the minimum NDVI for desert classes, $NDVI_\infty$ is the maximum NDVI for evergreen forest classes, and NDVI is the current NDVI value for the cell. f_g is constrained to the interval 0-1.

3.4.3 Archived Algorithm Output

The VIIRS Quarterly Surface Types IP, the VIIRS Surface Types-Olson IP, and the VIIRS Surface Types-Biomes IP will need to be archived for successive three month periods until such time as new quarterly products can be produced. The maximum, minimum and amplitude in VI will also be stored for use in the green vegetation fraction calculations, and will also be updated each quarter. Training pixels correctly and incorrectly classified will be saved for future analyses and algorithm refinement. Areas with significant land cover changes will be flagged. Several quality flags describing product quality will also be included in the output.

3.4.4 Variance and Uncertainty Estimates

The following is a list of potential sources of error that may be expanded upon in the future through a sensitivity analysis:

- Errors from EDRs upstream (e.g. VI)
- Bidirectional Reflectance Distribution Function (BRDF) related errors
- Topography
- Shadows
- Sub-pixel clouds
- Atmospheric correction errors
- Undetected land cover changes in training data

- Training data mis-labeling
- Mis-registration
- Sensor Drift and Calibration
- End members in mixture models
- Mixed pixels

3.5 ALGORITHM SENSITIVITY STUDIES

It is extremely difficult to accurately simulate global, temporally composited monthly spectral or thermal data, even with the relatively simple configuration of the AVHRR. This would imply simulating not only the spectral variability of cover types but also being able to realistically simulate the temporal evolution and temporal variability of these cover types globally. Such a simulation would test not only the current scientific understanding of coarse scale surface type patterns but also the current computational capabilities of many agencies and/or research groups. Nonetheless, that work is currently ongoing.

In sections 3.5.1 through 3.5.5 we present simulations using single date TM scenes and only the spectral information to classify various surface types. These sections are not intended to provide an error budget for this EDR because none can be realistically produced as previously described. However, these sections are useful to better understand the contributions of the principal sources of error to the selected classifiers, and can give indications of the relative magnitude or importance of each error term in the classification process. The sections also illustrate how the different classifiers respond to different kinds of error sources and thus may be used as an indicator of their potential robustness for future operational applications.

3.5.1 Radiometric Noise and Stability

The Quarterly Surface Type IP will be produced from time-composited data that will further reduce any important variability due to sensor noise (Holben, 1986). In any case, it is expected that the natural variability in the reflectance of each cover type at a coarse scale should be much greater than that introduced by sensor noise. Moreover, the decision tree is not particularly affected by noisy data because each pixel is classified and then noisy branches of the tree can be identified and pruned (Quinlan, 1993). Neural networks, however, can be affected by noisy data, but only from levels much higher than most normal sensor noise.

We performed the following procedure to illustrate the potential effect of random instrument noise on regional scale land cover classifications. We introduced sensor noise to simulated VIIRS radiances for 10 of the 21 IPO land cover classes from a Washington, D.C. TM scene. We generated spectral signatures for 10 VIIRS bands in the visible, near and mid-infrared for each cover type and introduced 1 percent random noise to simulate a nearly “ideal” case for land cover classification. We have then selected an area of 256 by 256 pixels to train our classifiers and tested these on a 512 by 512 pixel area surrounding the smaller area. In this case, the radiances for each land cover class tend to cluster together in spectral space and are easily distinguishable. Recognition accuracies for the C5.0 decision tree (DT) and the Learning Vector Quantization (LVQ) neural network classifier are nearly 100 percent, as expected.

We then introduced sensor noise according to seven different sensor noise models provided by Santa Barbara Remote Sensing (SBRS) (RIM Y1629). The same procedure as above was followed for both training and testing the classifiers. For both the DT and LVQ algorithms, the recognition accuracies for all sensor noise models up to 6 is greater than 98 percent. Even for sensor model 7, the accuracies are 81.8 percent and 76.4 percent for the DT and LVQ, respectively. All of these are above our threshold accuracy of 70 percent and approach or surpass our objective accuracy of 80 percent. Additionally, the DT algorithm was tested with a 10-fold cross-validation approach (Friedl and Brodley 1997; Quinlan, 1993) whereby 10 percent of the 512 by 512 area was randomly selected to produce 10 decision trees and subsequently tested on the remaining 90 percent of data. Again, the average accuracies are nearly 99 percent for all sensor noise models, with an average accuracy of nearly 83 percent for sensor model 7.

Note that we have tested this effect on a limited number of land cover classes, and have only introduced a small amount of variability into the radiances. In reality, natural land cover types will exhibit a much larger variability in radiances and reflectances. Some may be difficult to separate because of similarities in their spectral signatures. It is thus unlikely that we would obtain threshold accuracies with sensor model 7 in these situations. However, this analysis re-iterates that sensor noise appears to have a very small effect on the selected land cover algorithms. If the sensor is built according to requirements for Noise Equivalent Delta Radiance (NedL) from the VI and the Surface Reflectance IP, it should also satisfy those for the Surface Type EDR.

Because this EDR depends on temporal information, the accurate calibration and long-term stability of the sensor are important issues. However, techniques currently exist to correct the sensor drift in the bands of the AVHRR to provide internally consistent global data (Holben *et al.,* 1990; Los *et al.,* 1994). It is expected that these techniques could be tailored to VIIRS in the eventuality of any sensor drift. Also, these techniques have been successfully applied to generate a global NDVI data set from three different AVHRR sensors (Los *et al.,* 1994). Again, calibration and long-term stability effects will be more significant on the VI EDR than the land cover EDR (Los *et al.,* 1994; DeFries and Townshend, 1994b).

3.5.2 Band-to-Band Registration

A misregistration between spectral bands on different instrument focal planes should principally affect the regional land cover classification. This classification will use instantaneous spectral information, principally from VIIRS bands M4, I1, I2, I3, and M11. A sub-pixel shift in any of these bands may also introduce some variability to the training data. Thus, it could affect the overall classification accuracy, particularly in heterogeneous areas. This shift, however, would need to be on the order of 0.2 pixels or greater to have a significant effect (Townshend *et al.,* 1992).

For the band-to-band registration analysis we used the same simulated TM scene for the Washington, D.C. area used in Section 3.5.1. We shifted VIIRS bands I3 and M11 with respect to bands M1 through I2 to simulate a misregistration between spectral bands in different focal planes. The shifting was done in both the x and y directions for 2, 4, and 6 TM pixels of 30 m. This simulated a misregistration of approximately 15, 31, and 46 percent, respectively, of a nominal 390 m VIIRS pixel. After shifting, the spectral data, as well as the classification product, were degraded to 390 m spatial resolution by averaging 13 TM pixels. The subsequent

files were used in the C5.0 software with a 10-fold cross-validation approach. This means using 10 percent of the data for training, and testing the remaining 90 percent for 10 different decision trees. The original mean accuracy for the tree, produced from the unshifted data, was 87.1 percent. The accuracies for the data which were shifted 2, 4, and 6 pixels were, respectively 85.5 percent, 85.5 percent, and 85.4 percent, showing that, as expected, the decision tree algorithm is fairly "resistant" to band-to-band misregistration. It would be expected that the LVQ neural network would show similar results because they are both learning machine algorithms. This means that the algorithms learn patterns from training data, no matter how noisy or misregistered, and attempt to classify unseen cases based on these previously defined patterns. This also implies it would be very difficult to tell whether spectral bands are misaligned from the results of either algorithm. However, based on this analysis, we feel that if the requirements for this parameter are met for the VI EDR, Surface Reflectance IP, and albedo EDR, they should also be met for the Surface Type EDR.

3.5.3 Mixed Pixels

Because of the intrinsic spectral similarity of some land cover types it is possible to still have large misclassifications errors with data from a perfect sensor, perfectly co-registered, and with little or no systematic errors. Likewise, the aggregation of many surface types to a coarser scale creates mixtures of materials whose combined spectra may be very similar and thus difficult to separate. We have examined these errors, which can be loosely termed "algorithm errors", by using a subscene of a Landsat TM scene provided by the IPO for the Olympic peninsula of Washington state in the northwest Unites States (Figure 8). As can be seen in Figure 8, this scene is quite heterogeneous and contains 10 of the 21 IPO surface types, including several forest types (a mixed forest class has been added), soil types, and a good portion of urban areas. Therefore, we believe this scene to be a rather challenging test for any classification algorithm.

For each land cover type in Figure 8, we have assigned a surface reflectance for each of the VIIRS bands based on pure spectra provided in the NPOESS simulation toolkit. The TM pixels were then aggregated to a simulated mid-swath VIIRS resolution of 1200m by assuming a perfect MTF function. Likewise, the land cover classification was averaged and the class with the highest fraction within the coarser cell was retained as the dominant class for the VIIRS pixel. By aggregating the data we hoped to mimic the inherent natural variability that exists in most landscapes and by selecting the dominant type for training we further challenged our algorithm. Typically, training data are selected which represent land cover types that are over 90% pure and in our case, the dominant type could have a fraction as low as 30%. For our simulations we randomly selected 75%, 50%, and 25% of these data to train the decision tree and then tested the results of the trees on the remainder of the samples. For all simulations, 10 different runs were performed.

For the 75% training sample, the mean typing error for the 10 runs was 8.23% and 16.52% for the separate testing sample, showing that, as expected, this is a dominant source of error for this algorithm. For the 50% samples, the errors were 8.63% and 17.03% for the training and testing sample, respectively. Finally, the 25% samples produced errors of 9.38% and 18.87% for training and testing. Using the boosting option in C5.0, the training error were nearly removed for the 25% samples (0.3% error) and were only slightly reduced for the testing sample (15.0%). It is interesting to note that for all analyses, over 50% of the total error generated was from

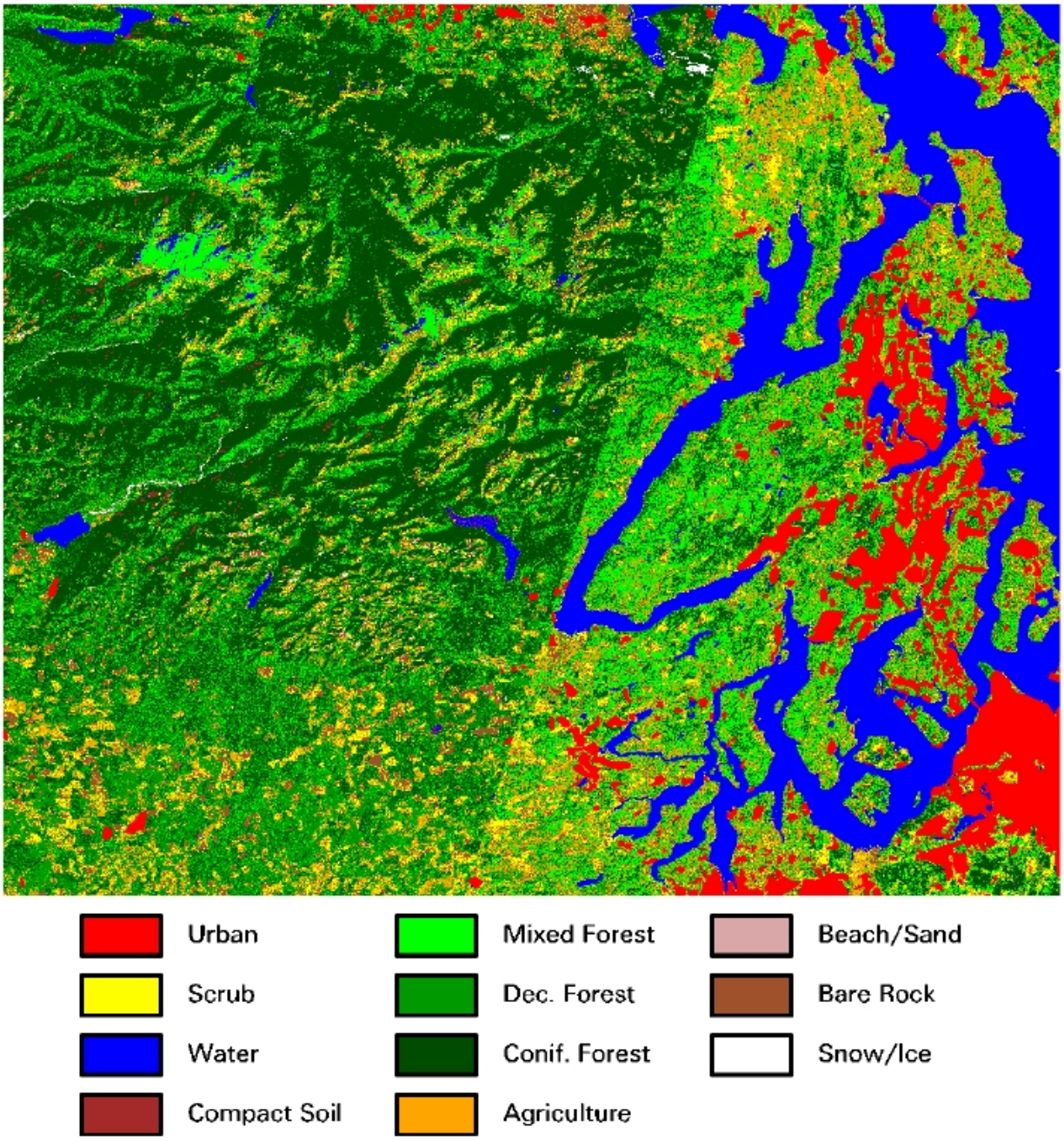


Figure 8. Subscene of Land cover classification of Olympic Peninsula Landsat TM test scene provided by the IPO.

confusion between the three forest types and we could expect simpler scenes to have much higher typing accuracies than this one. Nonetheless, these analyses show that natural spectral variability within a pixel is indeed a significant error source for classifications, using over 50% of the allowable error in the EDR. The analyses also imply that the use of the dominant class for training, which simulates the use of a previous global surface type database, may introduce errors that would need to be reduced by further screening the training data.

3.5.4 Misregistration of Training Data and Surface Reflectance

We have also simulated a potential misregistration of 1, ½, and ¼ VIIRS pixels between the training data and the VIIRS surface reflectances, corresponding to misregistrations of 1200 m, 600 m, and 300 m, respectively. The TOA BTs were not included in these analyses. Three shifts were performed on the data, first shifting the reflectances in the X direction with respect to the training data, then in the Y direction, and then in both the X and Y directions, which was a worse case scenario. Only the 25% training and testing samples were used in these analyses. The same data used in 3.5.3 were also used here and the aggregation error was included because the pixels were averaged after they were shifted. The results are summarized in Table 4.

Table 4. Typing errors due to misregistration of training data. Values given are the average of 10 different runs using 25% of the data for training and a separate 25% for testing. The second row of values for each shift are produced when using the boosting option in C5.0

1200 m Shift					
X Direction		Y Direction		X and Y Directions	
Train	Test	Train	Test	Train	Test
17.85%	39.54%	16.65%	36.67%	19.83%	44.20%
0.5%	37.0%	0.6%	34.2%	1.5%	39.7%
600m Shift					
13.43%	30.68%	13.10%	29.00%	15.72%	35.15%
0.90%	28.60%	0.50%	27.60%	0.90%	32.20%
300 m Shift					
10.49%	25.34%	10.95%	24.20%	12.25%	27.31%
0.10%	20.50%	0.30%	21.90%	0.80%	24.00%

As can be seen, the combination of pixel aggregation with either a 1200 m or 600 m misregistration in any direction causes the typing accuracies for this case to be below the threshold typing accuracy for this EDR. The boosting option only remedies this situation slightly in most cases and appears again to remove most of the errors in the training data. The typing errors are nearly doubled for the 1 pixel shift, while they increase between 10 to 16% for the half pixel shift when compared to the analyses in 3.5.3. The 300 m error adds between 5 to 8.5% additional error to the errors of the previous analyses. These same increases in errors are also seen in the test results when using the boosting option.

Clearly, a registration error of 300 m or less is necessary to produce acceptable results in the regional product and even in this case, the typing accuracy does not provide much margin should other sources of error be included. While the simulations presented here do represent a fairly stressing case for the regional product, it may not be uncommon to find this particular combination of land cover types on a particular VIIRS orbit. This 300m error, however, is larger

than what has been discussed in the past for land cover change detection (Townshend *et al.* 1992) and is currently proposed for MODIS. We fully expect that the VIIRS system will be able to do better than a 300-m registration error for the relevant bands.

3.5.5 Errors in Surface Reflectance

The VIIRS Quarterly Surface Type IP relies on surface reflectances which have been corrected for atmospheric scattering and gaseous absorption. This process of atmospheric correction nonetheless introduces some systematic errors into the input data that can potentially reduce the correct typing probabilities. We have simulated a random error in retrieved surface reflectances and examined its effect on the typing accuracies of the decision tree. Two types of analyses that perturbed the input spectra with random noise in reflectance were performed. In one, we added random noise as a percentage of the input reflectance, introducing noise as 10, 25, and 50% of the input value. In the second, we added the error as an absolute value in reflectance, with tests being performed with error values of 0.01, 0.02, 0.03, 0.04, and 0.05 in reflectance being introduced to the NPOESS toolkit spectra. The same Landsat TM scene used in Sections 3.5.3 and 3.5.4 was also used here.

10-fold cross validations were performed for each random error analysis, yielding typing errors of 1.2, 14.1, and 30.1 % on the test data for the 10, 25, and 50% random error cases, respectively. It should be noted that a 50% random error in surface reflectance introduces small relative errors for surface types with low surface reflectances (e.g. vegetation in the visible wavelengths) but very large errors for surface types with high reflectance values (e.g. vegetation in the near-infrared, snow, sand). We found that the majority of the errors introduced in these tests were caused by confusions between the different forest types as well as the different soil types.

The typing errors for the second set of analyses were 0.5, 6.2, 12.5, 19.5, and 26.1% for the 0.01, 0.02, 0.03, 0.04, and 0.05 absolute random errors in reflectance. Here the trends observed were opposite to those seen in the previous analysis, with low surface reflectances being affected much more strongly than high. For example, a 0.05 random error in reflectance could represent a very small error for snow (on the order of 10%), while the same value produced a coefficient of variation of over 200% for the needleleaf forest class in the red portion of the spectrum. Clearly, analyses which consider errors in the individual bands used may be needed in the future, but these results show the magnitude of the errors that can be tolerated by the algorithm and these are shown to be quite large in many cases. We expect errors in atmospheric correction for VIIRS to be on the order of 0.01 to 0.02 absolute.

As a final test, we have merged some of the smaller classes together to produce a new image with seven land cover classes (Table 5). Here we merged the mixed forest class with the deciduous forest class and all the bare classes (i.e., compact soil, beach/sand, bare rock, snow/ice) into one bare soil category. This new simulated classification retained the heterogeneity of the original scene and removed some of the smaller classes present. 10-fold cross-validations with the same absolute random error as before were performed. Here, the typing errors were 5.6% and 11.3% for 0.04 and 0.05 absolute error in reflectance, compared to 19.5% and 26.1% for the previous analyses, showing the effect of some of the cover types previously included. Based on all these analyses, we find that an absolute error of 0.04 in surface reflectance appears to be appropriate for the baseline approach to retrieve the Quarterly Surface

Types IP from VIIRS data. Analyses which incorporate the combined effect of the error sources studied, as well as other simulated scenes, may be included in future versions of this document.

Table 5. Land cover classes included in the surface reflectance error analyses and their proportion within the scene.

Class No.	Surface Type	% of Scene
1)	Urban	6.30
2)	Scrub	5.14
3)	Water	12.75
4)	Bare Soil	6.54
5)	Deciduous Forest	28.47
6)	Coniferous Forest	34.15
7)	Cropland	6.65

3.6 PRACTICAL CONSIDERATIONS

3.6.1 Numerical Computation Considerations

End-to-end tests on the required computer processing speed and data storage requirements will be performed as this algorithm matures.

3.6.2 Programming and Procedural Considerations

See Section 3.6.1.

3.6.3 Configuration of Retrievals

The Quarterly Surface Type EDR will be a Level 3 product in the Goode's Interrupted Homolosine Equal Area projection. Allowing for quality flags, the class numbers will be represented by 8-bit numbers whose value and order will be determined at a later date. The percentage of each cover type will be provided as a scaled 8-bit number as will the percentage of green vegetation fraction.

3.6.4 Quality Assessment and Diagnostics

Several quality flags will be produced from this EDR and will be arranged according to the following criteria:

- Red flags:
  - Missing/No data.
  - Ocean.
  - Clouds.
  - Inherited red flags from the upstream algorithms.

  - Solar Zenith Angle (SZA) threshold exceeds the no-report limit.
  - View Zenith Angle (VZA) threshold exceeds the no-report limit.

- Yellow flags:
  - Pixel filled during binning process.
  - Mixed Cloud/No Cloud.
  - Cloud/Topographic Shadows.
  - Solar Zenith Angle (SZA) threshold exceeds the threshold limit.
  - View Zenith Angle (VZA) threshold exceeds the threshold limit.
  - Inherited yellow flags from the upstream algorithms.
  - Active Fire in cell.
  - Active Fires has occurred in previous quarter.
  - Land cover change detected.
  - Pixel flagged as changed.
  - Training pixel misclassified.

- Green flags:
  - Training pixel (classified correctly).

Further development of these flags and fusion with the Land Quality Flag (LQF) structure described in *[V-4]* will be documented in Version 5 of this ATBD.

3.6.5 Exception Handling

This EDR will be produced for all land surfaces, including inland water bodies.

3.7 ALGORITHM VALIDATION

The validation of remotely sensed global land cover classifications presents a significant challenge to the land cover community. Not only are there no precedents, but there is no current accepted "state-of-the-art" validation technique. The validation of *operational* land cover classifications presents yet other challenges beyond those already present. However, by the time the NPOESS platforms are put into operation, the experience and knowledge in validating coarse scale land cover products will have been significantly advanced by current and planned global land cover classification activities. Principal among these activities are the validation of EROS Data Center's International Geosphere Biosphere Program Data and Information Systems (IGBP-DIS) 1 km global product, and the land cover products from MODIS. It is expected that the validation of the VIIRS surface type product would use data from validation sites created as a part of the above activities. The validation of the VIIRS surface type product is also expected to be performed in conjunction with the validation activities of other VIIRS products, particularly those of the Land IPT. This section briefly describes the objectives, methodology, and data requirements for the pre- and post-launch validation of the Surface Type EDR (40.6.4). It is expected to evolve throughout the lifetime of the project.

3.7.1 Objectives

The principal objectives of the validation are to verify that the global and regional land cover products generated from the input VIIRS data can achieve threshold requirements as specified by the IPO, and to provide a quantitative estimate of the recognition accuracy of the algorithm. For the Surface Type EDR, this implies that for either regional or global products, the dominant surface type identified by the algorithm must agree with the dominant land cover type at the surface at least 70 percent of the time. Secondary objectives are to identify both errors of omission and commission; to pinpoint error sources; to identify sources of confusion between surface types; and define avenues for algorithm improvements.

3.7.2 Pre-Launch Algorithm Validation

3.7.2.1 Methods

The Surface Type products will be validated by the following methods:

- Comparison to AVHRR 8 km and 1 km training data from the University of Maryland.
- Comparison to AVHRR 1 km training data produced from the IGBP-DIS global classification.
- Comparison to AVHRR 8 km and 1 km classification product from the University of Maryland.
- Comparison to IGBP-DIS 1 km classification product and degraded 4 km classification product.
- Comparison to SeaWiFS, POLDER or MODIS global classifications if available.
- Incorporation of high resolution TM land cover classifications.
- Comparisons against existing high resolution land cover databases (Global Land Cover Test Sites, Field campaigns, regional data sets).
- Regional verification by consultation with regional experts.
- *In situ* data from USDA-Agricultural Research Service MODIS validation activities.
- Scene simulations with all IPO scenes.
- Comparison of FAO soil map with STATSGO soil map over North America.

3.7.2.2 Data Needs

- AVHRR 8 km and 1 km global NDVI, Ch. 1-2 Top-of-Atmosphere (TOA) reflectances, Ch. 3-5 TOA brightness temperatures.
- For global product yearly metrics for maximum, minimum, mean, and amplitude in above data.
- UMD 8 km and 1 km classification.
- UMD 8 km and 1 km training data.
- UMD MSS scenes used for training if possible.

- IGBP-DIS classification and 12 month of maximum NDVI composites.
- Updated versions of all or any of the above.
- IGBP-DIS validation results when available.
- Thematic Mapper classifications.
- TOA reflectances or DN for above classifications when available.
- Thematic Mapper data for snow/land/urban discrimination.
- For IPO scenes, BOA reflectances and land cover type, at both full and degraded resolutions, realistic seasonal changes in reflectances, "natural" variability included if possible.
- Realistic top of canopy spectra for different land cover types.

3.7.2.3 Data Availability

We have acquired most of the data products listed in 3.7.2.2, and are already using them for preliminary verification of the selected algorithms. We will acquire the UMD 1 km training data and metrics in the near future. The IPO scenes will provide a good testbed for this EDR, because any combination of cover types can be simulated. Many other supporting data sets have been identified worldwide and will be acquired according to their criticality. These include the Global Land Cover Test Sites Project, the North America Landscape Characterization (NALC) project, and the Multi-Resolution Land Characteristics (MRLC) data set, among others. Global satellite data from SeaWiFS, POLDER, and MODIS may not be available before PDR, but will not be critical to validation, because our substantial holdings of AVHRR data are sufficient on their own. The purchase of additional Thematic Mapper data encompassing important IPO surface types may be warranted in the future, and will be coordinated with the Snow/Ice team.

3.7.2.4 Data Quality

Both the AVHRR 8 km and 1 km data contain processing artifacts and residual errors due to misregistration, BRDF, insufficient atmospheric correction, for example. These will affect the accuracy of the retrieval, but VIIRS data should improve performance. We already achieve threshold accuracies in all cases tested so far with AVHRR data.

Training data will change over time, particularly due to anthropogenic activities and inter-annual climate variability. Therefore, training data will need to be closely scrutinized to ensure that erroneous data are excluded. This will be done on a regular basis by either selecting the areas where both the neural network and the decision tree agree for training, or by running a cross-validation with the decision tree, using only those sites that are correctly classified in all tests.

Neither the FAO digital soil map of the world nor the STATSGO soil map of North America have been validated; this could be a problem for a general accuracy assessment. But because both these maps have been compiled from extensive field surveys, they are expected to be fairly accurate. We will need quantitative estimates of the accuracy of these maps.

Ideally, surface type information collected *in situ* is thought to be robust for validation of remotely sensed data. However, the point data need to represent the entire satellite pixel, which can be at the scale of one or several kilometers. The validation of a dominant surface type or the

percentage of cover types at satellite scale, from point data collected by human analysts, is thus influenced by scaling issues that will need to be addressed, most likely through statistical analyses and/or sampling strategies.

In most cases, land cover classifications at the scale of TM are not fully validated. In addition, the land cover classes used in these TM classifications are not easily translatable to those required by the IPO. Also, if a TM classification is 80 percent correct, and our algorithm agrees with the TM data 80 percent, the total accuracy could be as low as 64 percent. Clearly, these are issues that will need to be further explored. We anticipate that the IGBP-DIS and MODLAND activities will develop well validated products in the future.

Although it is difficult to simulate "natural" variability in the spectral reflectances of surface types in the IPO scenes, these provide an excellent testbed for error budget analysis that cannot be duplicated from AVHRR data. We expect the principal sources of error to be principally related to misregistration, atmospheric correction, BRDF correction, sub-pixel clouds, and, to a smaller extent, sensor noise and band-to-band registration. Each of these error terms will be explicitly modeled using the IPO scenes and quantitative estimates provided.

3.7.3 Post-Launch Algorithm Validation

3.7.3.1 Methods

The methodology used for post-launch validation of the Surface Type EDR will be essentially the same as that used in the pre-launch era, because we can use previously generated coarse- and high-resolution land cover classifications and data acquired over the MODLAND and IGBP-DIS validation sites. However, because of the operational nature of the VIIRS EDRs, the post-launch validation must also include simultaneous data acquisition from VIIRS and other satellite platforms, and near- or at-surface validation observations. This can be accomplished either by participating in multi-disciplinary field campaigns, or by organizing field campaigns just for VIIRS or NPOESS validation. These field campaigns would be much narrower in scope than field campaigns such as FIFE, HAPEX-Sahel, and BOREAS. For the surface type EDR, the simultaneous acquisition of VIIRS data together with satellite, airborne, and surface data would not be critical, because surface types rarely change much over small time periods. However, point data would be needed to validate high spatial resolution land cover maps from TM or ETM+, for example, and these in turn would be used to validate the coarse resolution VIIRS products. It would also be desirable to organize field campaigns where representative samples for the 21 IPO surface types could be found, should those not already be available within the MODLAND validation sites.

3.7.3.2 Data Needs

- Updated 1 km training data from UMD.
- Training data from MODLAND and/or IGBP-DIS validation sites.
- Validated global land cover classifications from MODIS.
- Validated global land cover classifications from other sensors (AVHRR, Vegetation, ASTER).
- Regional and/or local validated land cover databases.

- Land cover change product from MODIS.
- Validated high spatial resolution land cover classifications.
- In conjunction with field campaigns, high spatial resolution classifications validated from field observations, and acquired within two weeks of VIIRS acquisition.
- *In situ* data from USDA-Agricultural Research Service MODIS validation activities.
- Updated soil maps if available.

3.7.3.3 Data Availability

We do not yet know which data will be available to us for our validation exercises. Concurrent MODIS and ETM+ data would aid validation immensely, but those sensors may not be in operation in the NPOESS post-launch era. This reinforces the need for participation in and/or organization of field campaigns. Most of the MODLAND sites are likely to still be in operation; even if they are not, the appropriate data can be acquired and verified through personal contacts. Data from other sensors that have yet to be launched may also be available.

3.7.3.4 Data Quality

The issues of data quality presented in 3.7.2.4 are likely to also be applicable in the post-launch era of NPOESS. However, we expect our pre-launch simulation and validation activities to provide us the answers to many questions, particularly those related to error sources and scaling issues. We also expect that the data quality of the data used for post-launch validation will be significantly superior to current products, and that much will have been learned from the MODLAND validation activities. Because data sources available for validation are likely to be quite good yet numerous, the issue will not necessarily be one of data quality, but of data *quantity*.

4.0 ASSUMPTIONS AND LIMITATIONS

4.1 ASSUMPTIONS

- A major assumption is that the Earth's land can be subdivided into *discrete* land cover classes or surface.
- A principal assumption of a supervised classification is that the training data are representative of global or regional patterns.
- The assumption is also made that land cover types exhibit fairly unique spectral and/or temporal properties and thus can be successfully separated.
- Training data have not changed composition since they were delineated and if they have this change can be flagged.
- We assume that most of the input data from other EDRs will be provided at, or better than, threshold accuracy.

4.2 LIMITATIONS

- Urban areas are difficult to classify because of their inherent variability at the global scale. We may be able to use city lights data from low light sensor to produce the Urban layer.
- Wetlands will be difficult to separate because of spectral similarities to other IGBP types.
- The direct retrieval of the soil classes specified in the SRD from VIIRS optical/thermal data may not be achievable considering the high degree of variability in soil reflectances, and other confounding factors.
- Some Grasses and cereal/crops will be confused.
- Some Rice paddies and wetlands will be confused.

5.0 PRELIMINARY RESULTS

5.1 PRELIMINARY TESTING OF CLASSIFIERS

5.1.1 Maximum Likelihood Classifier

This approach was tested to provide a baseline for comparisons with other algorithms that are expected to perform significantly better (DeFries *et al*., 1995a; Friedl and Brodley, 1997; Weiss and Kulikowski, 1991). We believe that an MLC, if appropriately parameterized, can approach threshold accuracy but its assumption of a normal distribution is its greatest disadvantage, particularly at a global scale. This approach also serves to train the analyst in more advanced features of the ERDAS image processing software, how to produce map outputs of classified products with this software as well as acting as a data exploration tool.

A first test was carried out with 24 AVHRR metrics for 1984 (i.e., minimum, maximum, mean and amplitude for NDVI, CH1, CH2, CH3, CH4, CH5) and the training data set from the University of Maryland (UMD) (DeFries *et al*., 1998a). The first task was to create signatures for each of the training sites. In this particular case, because the mean and amplitude of each metric is produced from the minimum and maximum values, the covariance matrices are not invertible and thus it is impossible to perform an MLC with these data. Several combinations of metrics were tested but in each case the covariance matrices for one or more cover types were not invertible. It was thus decided to focus on the minimum and maximum values for each of the metrics.

An examination of these data revealed that the metrics for channels 3, 4, and 5 are correlated and may not provide much more additional information beyond that provided by one single band. In a second test the minimum and maximum for NDVI, CH1, CH2 and CH5 were input into an MLC. Table 6 shows the percent of pixels that were classified accurately when compared to the training data.

Table 6. Per class accuracies for MLC using training data from DeFries *et al.* (1998a).

	[Code]	Surface Class (UMD)	MLC Accuracy (%)
1)	EBrF	Evergreen Broadleaf Forest	93.8
2)	ENeF	Evergreen Needleleaf Forest	68.2
3)	DNeF	Deciduous Needleleaf Forest	92.9
4)	MixF	Mixed Forest	69.4
5)	Moss	Mosses and Lichens	95.8
6)	Shrb	Open Shrubland	88.7
7)	Bare	Bare	98.3
8)	DBrF	Deciduous Broadleaf Forest	77.4
9)	Crop	Cropland	81.1
10)	Gras	Grasslands	63.6
11)	WdGr	Wooded Grasslands	90.4
12)	Wdls	Woodlands	43.6
13)	Bush	Closed Bushland/Shrubland	80.3

Overall, 7443 pixels were classified accurately out of a possible 9306 training pixels for an overall accuracy of 79.98 percent, which is close to our original objective accuracy. The mean accuracy for the 13 classes was 80.26 percent, slightly above the objective accuracy. These results are indeed very encouraging but they should be considered objectively. A qualitative comparison of the MLC output with that of DeFries *et al*. (1998a) shows several expected problems, particularly in the accurate representation of grasslands, croplands and woodlands. However, some cover classes such as evergreen broadleaf forests and bare surfaces appear to be fairly accurately represented. In the cases with low accuracies, typically the distributions of land cover information are multi-modal and overlap, thus making it difficult for the MLC to classify accurately. In these cases it may be possible to split two or more uni-modal clusters from the training data and then use these in an MLC. In any case, these results identify the problematic land cover classes on which we may need to focus more of our energies in order to increase classification accuracies. This test also shows that even fairly simple approaches can yield some promising results.

5.1.2 Decision Tree Classifier

The C5.0 decision tree software (Quinlan, 1993) was also tested with the 24 AVHRR metrics and the training data from UMD. In this program there are several options which can be used and which yield different levels of accuracy. Many different options were tested, but the principal test was carried out by using 60 percent of the training data to produce a decision tree and the remaining 40 percent of the data for testing results. The confusion matrices are shown in Figures 9a and 9b for the same land cover classes listed in Table 6.

Decision Tree		Rule Sets	
Number of Nodes	Errors	Number of Rules	Errors
225	112 (2.0%)	161	157 (2.8%)

5584 cases	Training Data Classified as:												
Code	1 EBrF	2 EneF	3 DNeF	4 MixF	5 Moss	6 Shrb	7 Bare	8 DBrF	9 Crop	10 Gras	11 WdGr	12 Wdls	13 Bush
1 EBrF	795	1						2	2				
2 ENeF	2	251		4				6	1	1	1		
3 DNeF		3	32	1									
4 MixF	2	6		385				4	2			1	
5 Moss		2			142							1	
6 Shrb						338			2	2			2
7 Bare							726		8				1
8 DBrF	1	1		6				384	3	1		1	
9 Crop								6	1106	9		3	
10 Gras						1	1	1	23	765	1	1	
11 WdGr									1	2	73		1
12 Wdls	3	1						1	7	3		262	
13 Bush						3			13	1	1		168

Figure 9a. Confusion matrices for surface types listed in Table 6. Matrix for 60 percent training sample.

Decision Tree		Rule Sets	
Number of Nodes	**Errors**	**Number of Rules**	**Errors**
225	380 (10.2%)	161	365 (9.8%)

3722 cases	Test Data Classified as:												
Code	1 EBrF	2 EneF	3 DNeF	4 MixF	5 Moss	6 Shrb	7 Bare	8 DBrF	9 Crop	10 Gras	11 WdGr	12 Wdls	13 Bush
1 EBrF	487	11		4				7		5		6	
2 ENeF	2	141	2	14	1			8	3	1		1	1
3 DNeF		8	21		1				3				
4 MixF	10	17		250	1			7	1	3			
5 Moss		1			92							1	
6 Shrb						231	1		2	5			3
7 Bare						3	461		2	5			
8 DBrF	8	1		8				209	7	1		10	1
9 Crop	4	3	1					11	695	18		2	1
10 Gras	2	1				12	2		41	473	3	3	3
11 WdGr						1			3	3	30		1
12 Wdls	8	1			1			2	14	10	6	165	
13 Bush		1				6			14	6	2	3	102

Figure 9b. Confusion matrices for surface types listed in Table 6. Matrix for 40 percent testing sample.

These results are again extremely encouraging in that the decision tree generated from 60 percent of the training data has almost 90 percent accuracy when used on the test data. One drawback from this approach is that there are 225 nodes for the tree (161 when using rulesets) which are difficult to implement on imagery in the ERDAS package. We consulted with the author of C5.0 to ascertain how this process can best be expedited and automated. We also expect that when these decisions or rules are implemented on actual imagery that we will find areas with obvious erroneous land cover types, as was found by the UMD team. These areas will be selected for further testing by identifying those nodes which contribute to their mis-classification and several tree-pruning methods will be examined.

A second test was performed on these same data but this time using an n-fold cross validation. In a cross-validation, the training data are divided into n random blocks from which decision trees or rulesets are produced and then successively tested on the remaining blocks of data. The results for a 10-fold cross validation are presented below, producing an average accuracy of 91.6 percent and 91.9 percent for decision trees and rulesets, respectively. Again, these methods produce a large number of decisions and rules but our future work will emphasize their optimization.

Table 7. Results for a 10-fold cross validation using the decision tree.

N-Fold	Decision Tree		Rules	
	Size	Errors (%)	Number	Errors (%)
0	279	8.4	199	8.3
1	285	8.8	201	8.2
2	294	9.7	207	9.0
3	290	8.5	194	7.6
4	297	7.8	204	7.6
5	285	7.6	206	7.4
6	279	8.2	203	7.8
7	285	8.5	196	8.7
8	294	9.1	185	9.5
9	302	7.6	195	7.2
Mean	289.0	8.4	199.0	8.1
SE	2.4	0.2	2.1	0.2

5.1.3 Neural Networks

A back-propagation neural network (NN) was tested on the same data used for the decision tree, but it failed to converge after 1000 iterations. Other input data may need to be input for this NN to converge.

The ARTMAP neural network was also tested by using the simulated TM scenes with 1 percent of noise added. To train the neural network, a number of pixels were randomly selected as the training data set, and a number of pixels were selected as the test data set. Two TM simulated scenes were used, one 256 by 256 in size, and the other 512 by 512 in size. The highest discrimination rate for the training data set is 79.1 percent and highest prediction rate is 79.4 percent. This result is not very encouraging because the simulated data have very low sensor noise and are basically the pure class spectra. Further investigation of other neural network algorithms as thus deemed necessary.

Overall, the above activities can be characterized as very successful, yielding results which surpass threshold and in some cases objective accuracies. It should be emphasized that these are, however, preliminary tests which do not contain the full complement of land cover classes required by the IPO.

5.2 IMPLEMENTATION OF DECISION TREE RULES

In the previous section we showed results from the C5.0 decision tree software using the 24 metrics of DeFries *et al*. (1998a). That particular tree had an accuracy of 98 percent on the training data and nearly 90 percent on test data but was very large (i.e., 225 nodes). A large decision tree is one drawback of the decision tree algorithm. It has a tendency to overfit the data and then it needs to be pruned back.

At this time the decision tree needs to be pruned and implemented by hand for input into the ERDAS software so the results for large trees are difficult to display for quick analysis. Two methods available in C5.0 were tested. One pre-prunes the tree according to specific parameters and the other specifies the minimum number of cases that can be passed down a branch. Of the several tests we performed one particular test with a minimum of 8 cases was found to be a good compromise between overall accuracy and the size of the tree. This test still had accuracies of 93.5 percent and 88.9 percent on the training and test data (Figures 10a and 10b), respectively, but only had 109 nodes.

We chose this tree for input into the ERDAS package. This was done after further pruning the tree by consideration of majority nodes and those with large numbers of misclassified pixels. The resulting decisions were input into ERDAS. These results are shown in Figure 11.

The results from this tree are quite good, agreeing very well with the 8 km classification of Defries *et al*. (1998a). However, similar problems to those encountered by those authors were also found on this classification. The good amount of open shrublands (yellow) present in the Sahara as well as the presence of deciduous needleleaf forest (cyan) in Alaska are two examples that, even though classification accuracies may be quite good, when rules are applied at the global scale some errors will still persist. This could be due to the fact that training areas are not sufficiently representative of the inherent natural variability or that the pruning methods are imperfect. Again, the refinement of our decision tree algorithm is the topic of ongoing research.

5.3 COMPARISON OF DECISION TREE AND LVQ ON 8 KM GLOBAL AVHRR DATA

We have performed a cross-comparison between the accuracies obtained by the decision tree and the Learning Vector Quantization (LVQ) Neural Network on common input data. We have randomly selected 10, 20, 30, 40, 50, 60, and 70 percent of the training data from DeFries *et al*. (1998a), used these to train the classifiers, and then tested the trained classifiers on the remaining blocks of data. The algorithm descriptions are provided in Section 3.4. Figure 12 summarizes the principal results from this comparison.

As can be seen, the accuracy of the algorithms increases with the number of pixels used both for the training and testing data sets. The performance of the DT is slightly better than that for LVQ for all cases, and is within the 80 to 90 percent accuracy obtained by DeFries *et al*. (1998a). All algorithm performance tests are well above our objective accuracy of 80 percent. It is interesting to note that particular land cover classes which have lower accuracies in one algorithm may have higher accuracies in the other. This is the case for croplands, for example, which have a higher accuracy in the LVQ when compared to the DT. The opposite is true for deciduous needleleaf forests. This indicates that a hybrid classifier as we have proposed may be appropriate because it will use the algorithms which are shown to perform optimally for each particular cover type.

Decision Tree		Rule Sets	
Number of Nodes	Errors	Number of Rules	Errors
109	364 (6.5%)	88	374 (6.7%)

5584 cases	Training Data Classified as:												
Code	1 EBrF	2 EneF	3 DNeF	4 MixF	5 Moss	6 Shrb	7 Bare	8 DBrF	9 Crop	10 Gras	11 WdGr	12 Wdls	13 Bush
1 EBrF	779	2		1				6	1			4	1
2 ENeF	17	232	1	12					10	1		3	
3 DNeF		1	38		1								
4 MixF	4	9	3	393				3	7				1
5 Moss		3			134				1			1	
6 Shrb						314			2	4			7
7 Bare						4	751			1			
8 DBrF	19	3		16				352				4	
9 Crop	3	1				4		1	1071			11	2
10 Gras	6			1		3	15	1	61	699		8	7
11 WdGr									2	2	56	1	5
12 Wdls	10				8			7	9	1		234	
13 Bush						6			14			4	157

Figure 10a. Results for C5.0 decision tree using 24 AVHRR metrics and 13 land cover classes. Matrix for 60 percent training sample.

Decision Tree		Rule Sets	
Number of Nodes	Errors	Number of Rules	Errors
109	414 (11.1%)	88	425 (11.4%)

3722 cases	Test Data Classified as:												
Code	1 EBrF	2 EneF	3 DNeF	4 MixF	5 Moss	6 Shrb	7 Bare	8 DBrF	9 Crop	10 Gras	11 WdGr	12 Wdls	13 Bush
1 EBrF	502	3		3				10	1	1		5	1
2 ENeF	8	123	3	13	1			1	10		2	3	
3 DNeF		2	21	3					3				
4 MixF	5	8	1	238				6	11				
5 Moss				2	97							1	
6 Shrb						221	7		3	14			12
7 Bare						2	446			1			1
8 DBrF	19	4		12				193	3			7	
9 Crop	5	2		2		6		4	695	18	4	11	3
10 Gras	6	1		2		9	11	4	42	450	1	8	2
11 WdGr									3	2	34	1	9
12 Wdls	7	1		1	10			5	10	7	2	172	
13 Bush						13		4	10		1	6	105

Figure 10b. Results for C5.0 decision tree using 24 AVHRR metrics and 13 land cover classes. Matrix for 40 percent testing sample.

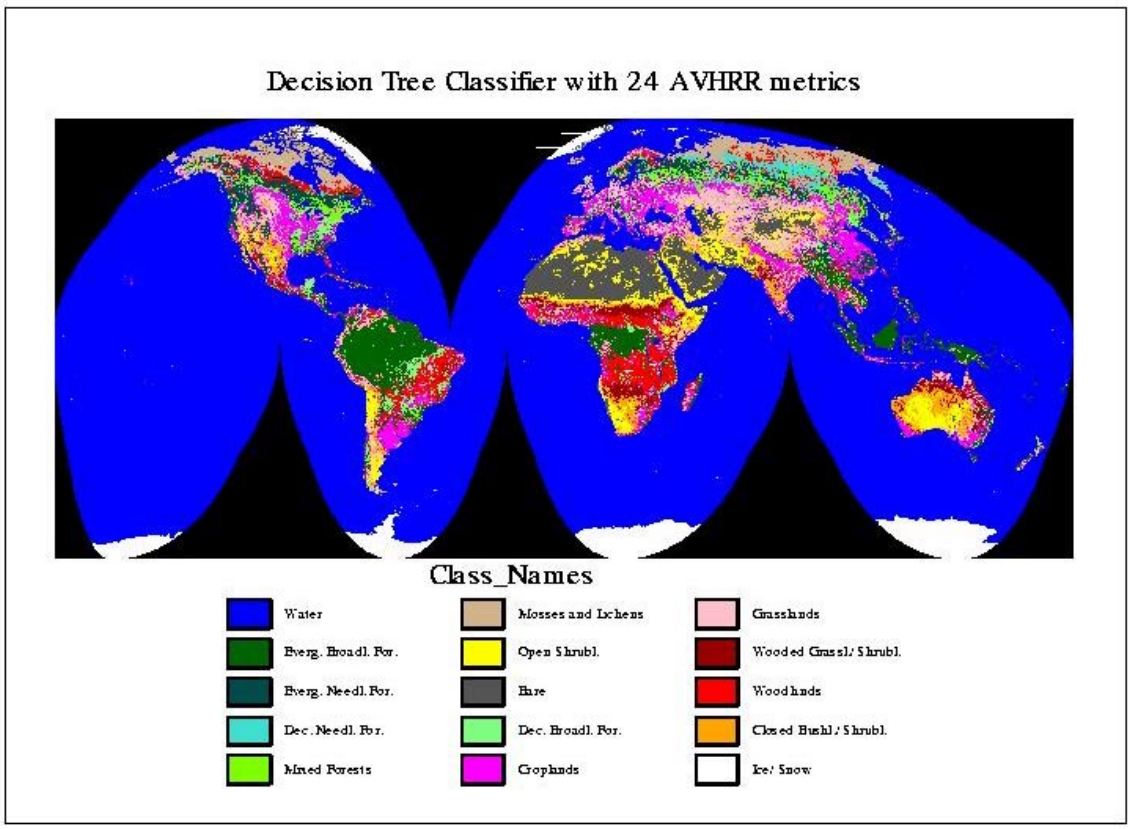


Figure 11. Decision rules from decision tree in Figure 9 implemented onto imagery.

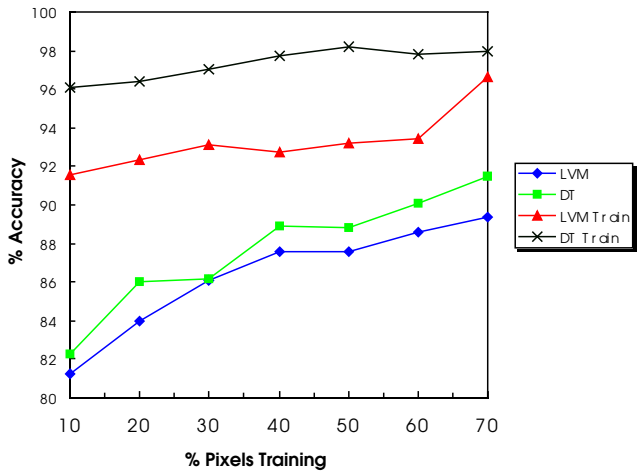


Figure 12. Percent accuracy of Decision Tree (DT) and Learning Vector neural network (labeled as LVM) as a function of the percentage of pixels used in training the algorithm. Accuracies for both training (i.e., DT Train) and testing data sets are shown.

5.4 DEVELOPMENT OF NEW 8 KM TRAINING DATA

While the training data supplied by the University of Maryland are very useful, the land cover classes that they represent lack some of the land cover classes specified by the IPO. We therefore needed to expand and supplement our training data. We chose EDC's 1 km AVHRR classification (Loveland and Belward, 1997) for this.

The first task was to re-code the EDC classes into an IPO/IGBP classification scheme, assuming only one bare surface class (desert). We took the most detailed of the EDC classification products, the Global Ecosystems Framework scheme with 99 land cover classes, and determined which of these classes were grouped into the EDC IGBP classification classes. By mapping the Mire, Bog, Fen and Marsh Wetland Global Ecosystems classes to the IPO class Marsh/Bog, and the Wooded Wet Swamp and Mangrove Global Ecosystems classes to the IPO class Swamp, we were able to split the permanent wetlands IGBP category into the appropriate IPO classes. This does not imply that they will be easily separable on the AVHRR data but that we do need to investigate their separability. This also makes the assumption, since we have no working definitions of the IPO land cover classes, that Marsh/Bog represents wetlands with primarily grass species while the Swamp category is dominated by woody species and trees. In the same fashion as with the wetlands classes we also used the Barren Tundra Global Ecosystem class to add a tundra class onto our modified IGBP scheme. Table 8 shows the correspondence between the IPO classification scheme, our modified IGBP classification scheme, and that used by the University of Maryland.

It should be noted that sometimes two UMD and IGBP classes are mapped to one IPO class and in other cases, some of the classes in the UMD and IGBP schemes do not have a corresponding class in the IPO scheme. The new scheme, nevertheless, does offer some added flexibility, no matter what the classification scheme that is ultimately chosen. Also, if we consider all the soil classes as one, and the flooded land class as a special case, we have now covered all the principal IPO classes and we are in a position to provide classifications with more land cover classes. This also means that the land cover classification scheme to be used for MODIS can be easily incorporated into our algorithm.

The next step in the process involves aggregating the 1 km classification to 8 km resolution so that we may be able to use it with our 24 AVHRR metrics (Table 9). For this, a program was used which outputs the percentage of each 1 km class within the coarser resolution pixel, as well as the dominant class within the 8 km pixel. It was hoped that pixels with 100 percent of the particular class at 8 km could then used to enlarge our existing training data set. However, the distribution of these pixels did not correspond very well with the percentage of each particular class in the 1 km data, with large classes such as Snow/Ice and Bare dominating at the expense of other classes. Thus, we adjusted the percentage of 8 km pixels for each class to more closely resemble those of the 1 km data. Table 9 shows the percentages that were used as thresholds and the resulting number of pixels mapped to each land cover class.

Table 8. Correspondence between IPO land cover classes, our modified IGBP system, and the University of Maryland's scheme.

IPO Classes	Modified IGBP Classes	UMD Classes
Tropical Forest	Evergreen Broadleaf Forest	Evergreen Broadleaf Forest
Coniferous Forest	Evergreen Needleleaf Forest	Evergreen Needleleaf Forest
Deciduous Forest	Deciduous Broadleaf Forest	Deciduous Broadleaf Forest
Savanna	Savanna, Woody Savanna	Woodland, Wooded/Grassland
Cropland	Cropland, Cropland/Natural Vegetation Mosaic	Cropland
Grassland	Grassland	Grasses
Brush/Scrub	Open Shrubland/Closed Shrubland	Closed Bushland/Shrubland, Open Shrubland
Tundra	Tundra (Added)	Mosses and Lichens
Swamp	Permanent Wetland	n/a
Marsh/Bog	Permanent Wetland	n/a
Desert	Bare/Sparse Vegetation	Bare
Rocky Fields	n/a	n/a
Gravel	n/a	n/a
Sandy Soil	n/a	n/a
Loam	n/a	n/a
Clay	n/a	n/a
Peat	n/a	n/a
Urban/Developed	Urban/Built-up	n/a
Flooded Land	n/a	n/a
Snow/Ice	Snow and Ice	Bare
Water	Water Bodies	Water
n/a	Deciduous Needleleaf Forest	Deciduous Needleleaf Forest
n/a	Mixed Forest	Mixed Forest

It can be seen that most training pixels are still quite pure for most classes, with greater than 57 out of the possible 64 1 km pixels being classified as the same class in the 8 km data. Some small classes such as swamp and urban/developed needed to be augmented quite a bit, but would still remain the dominant class at the 8 km scale. For the urban class, the EDC classification uses a vector overlay of the digital chart of the world and appears to provide good information in the United States and Europe, but does not show the large cities of China and India very well. In this case, purposely increasing the size of the area covered may be particularly appropriate. Additionally, training areas were delineated manually on the imagery to decrease the size of both the Snow/Ice and Bare classes. The resulting training data set now has 712,651 pixels for training (compared to 9,306 from UMD). It would be expected that this data set would be substantially noisier than the UMD data set, but it may provide a more realistic or more global training set. Figure 13 shows the spatial arrangement of the new training pixels on the globe.

Table 9. Percentages used to adjust number of training pixels for each land cover class in our modified IGBP/IPO scheme.

	Class Name	Percent Threshold (≥ %)	Number of Pixels
1	Evergreen Needleleaf Forest	96	27931
2	Evergreen Broadleaf Forest	100	55955
3	Deciduous Needleleaf Forest	96	11110
4	Deciduous Broadleaf Forest	90	14047
5	Mixed Forest	90	26574
6	Woody Savanna	90	49716
7	Savanna	96	51569
8	Grassland	96	57633
9	Closed Shrubland	90	12194
10	Open Shrubland	100	98224
11	Swamp	57	454
12	Marsh/Bog	82	6446
13	Tundra	94	13084
14	Cropland	90	74741
15	Cropland/Natural Vegetation Mosaic	90	44667
16	Urban/Developed	50	1276
17	Bare/Sparsely Vegetated		83278
18	Permanent Snow and Ice		83722
19	Water	UMD Water Mask Used	
	Total (water not included)		712651

For testing purposes, we included all of our new training pixels into a decision tree using the 24 AVHRR metrics from UMD. The software took over 113 minutes to run with the 712,651 samples and contained 42,120 nodes but still had an overall accuracy of 95.9 percent. As expected, classes such as Swamp, Marsh/Bog, Urban/Developed and Cropland/Natural Vegetation Mosaic had poorer accuracies than the others, and in particular, the Urban class had an accuracy below 50 percent. Because urban areas are composed of so many different components that can vary depending on geographical location, it may be extremely difficult to accurately classify this land cover type at the global scale. The MODIS ATBD does not give any clues as to how this cover type is separated in their algorithm. We may need to use a static database for this cover type as well.

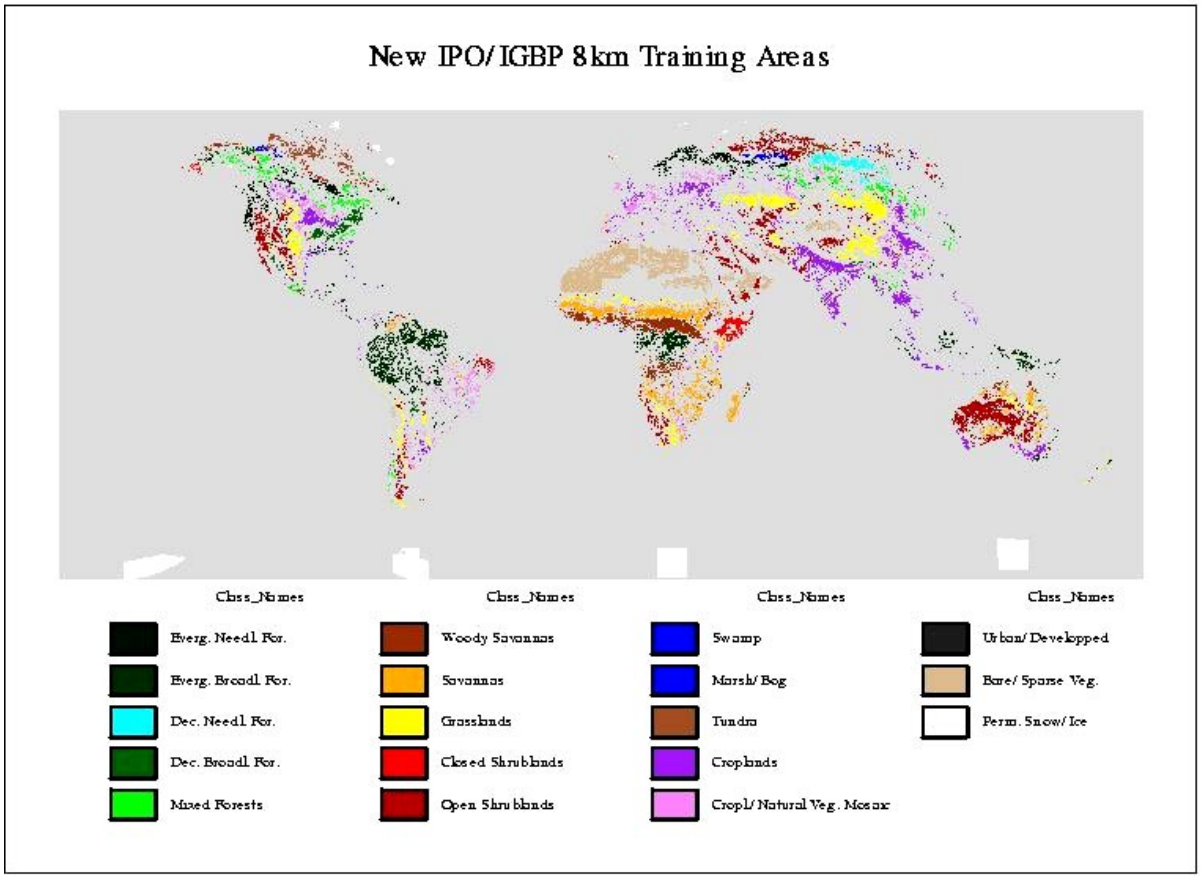


Figure 13. New 8 km training areas for IPO/IGBP land cover classes derived from EDC 1 km global classification.

Figure 14 shows results produced using these new training data over North America. To produce these results, several tests with different decision tree options were run and it was found that the boosting option in C5.0 provided the best results for this data set that used the 24 UMD metrics for 1984 coupled to the degraded EDC classification for 1992. With the boosting option, a series of 10 decision trees are built and each successive tree pays particular attention to any erroneous results from previously generated trees. Typically the boosted tree provides recognition accuracies which are 10 percent or greater than those from a normal tree but the process is, however, more computer intensive.

The Flooded land class is not considered here in Figure 14. Loam, Sandy Loam, Clay, Peat, Gravel, and Rocky Fields cover types are included in the Desert cover type. The training data were generated from relatively pure (greater than 90 percent for most classes) 8 km pixels. Using the boosting option, the recognition accuracy is over 99 percent for 70 percent randomly selected pixels in a) and 88.2 percent for the remaining 30 percent in Figure 14a. We have also produced a Dominant Class image (not shown) from the percentages of each cover type within the degraded 8 km pixels. For the entire North American continent, the overall recognition accuracy (i.e., against training, testing, and unseen cases) is approximately 74 percent. These results are extremely encouraging, considering that the EDC data are for 1992 and the AVHRR metrics are for 1984.

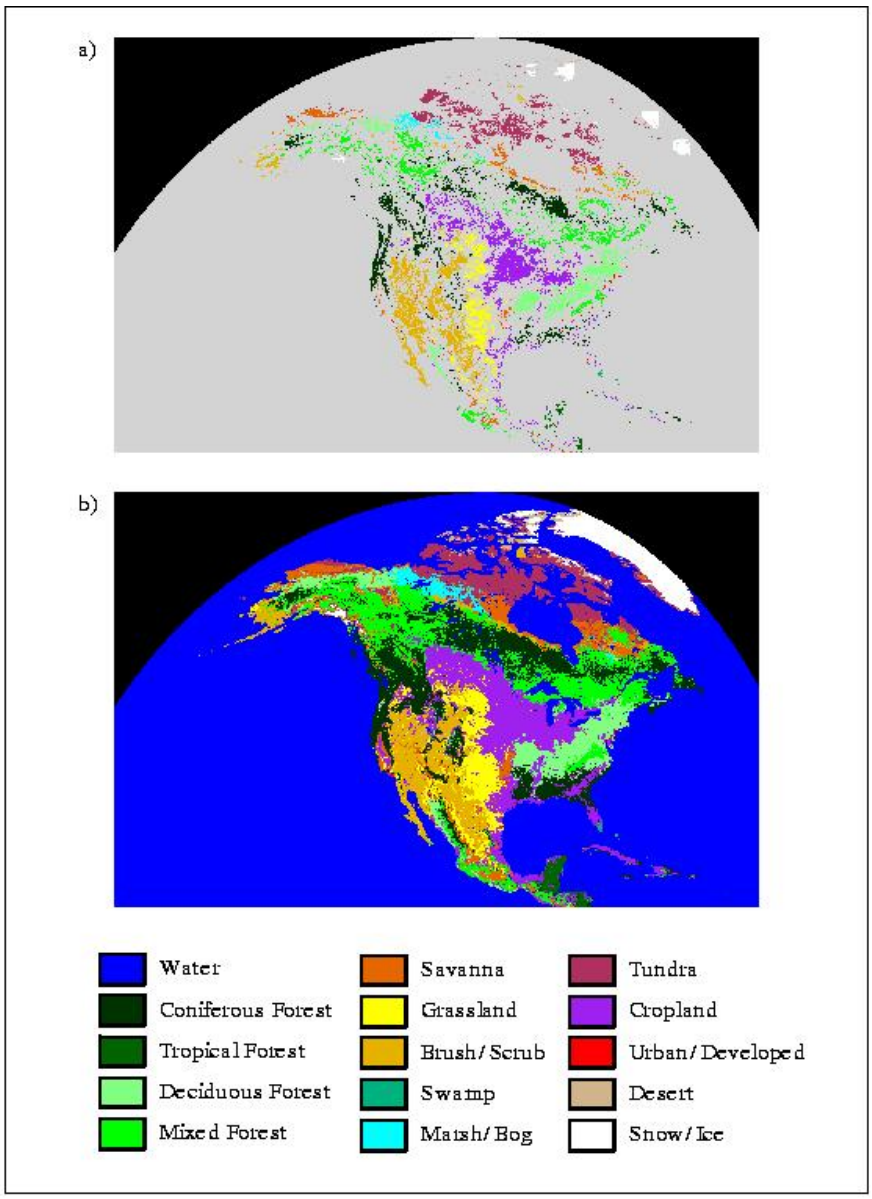


Figure 14. a) 8 km training data produced from degraded 1 km AVHRR global land cover classification from EROS Data Center (EDC) (Available from *http://edcwww.cr.usgs.gov/landaac/glcc/globe_int.html*). b) Output from C5.0 decision tree with boosting option (Quinlan 1993) using 24 AVHRR metrics for 14 of the 21 IPO cover types, plus a Mixed Forest type.

This is our first attempt at using an IPO classification scheme at the continental scale, albeit with only one bare soil class (Desert). In general, results are very encouraging but low individual class accuracies for land cover classes such as Urban/developed and Desert, for example, point to the need for either reducing the number of training samples or somehow cleaning up those that are included. Closer examination of the results and training data also revealed what may be errors in the EDC classification, such as wooded grasslands (mapped as Savannas in Figure 14) in Northern Alaska and the presence of shrublands on the Western portion of Baffin Island and on

Queen Elizabeth Island. Consultation of other Alaska vegetation maps show the same areas classified as principally tundra with shrubs. We will contact the EDC with these findings. Finally, this particular data set and subsequent results are quite unique to our knowledge and present several research problems that will also be addressed in the future.

5.5 DEVELOPMENT OF PROTOTYPE VIIRS 1 KM CLASSIFICATION

The only currently available data set for 1-km testing on a global scale is based on AVHRR, and it is difficult to either simulate or break down the related errors into individual sources, as previously discussed in section 3.0. However, because the input data are global, monthly composited AVHRR data, all of the AVHRR error sources are imbedded within the data sets and as such provide a much more stringent and realistic test of the algorithm performance as would be provided by simulated data.

The final step in the algorithm development for this EDR involved the testing of the decision tree software on the full resolution 1 km global AVHRR product from EDC. Because of computational limitations (a full version of the data set requires in excess of 60 GB of storage space), we were only able to test a sub-sample of the global data set by extracting every 5th pixel from the full resolution 1 km AVHRR data. The input metrics and training data used are described in detail in Hansen *et al.* (2000). We have extended this global training data set to the full IGBP classes. Specifically, we have re-labeled some of the training data from Hansen *et al*. (2000) into the IGBP wetlands class (e.g. Everglades, Louisiana Swamps). We have also delineated other wetlands training data on the AVHRR data directly in regions with known large wetlands areas (e.g. Pantanal in Brazil, Okavango Delta in Africa, Russian wetlands east of the Urals). Urban areas were also delineated in this fashion for large urban complexes all over the world. Snow and Ice training data were selected from large ice caps or glaciers in Greenland, Alaska, and Iceland. Finally, the IGBP Croplands/Natural Vegetation Mosaics class was modeled by randomly mixing from 10 to 60% of the croplands metrics with either forests, shrublands, and grasslands metrics, as specified by the IGBP definition for this class (see Table 3). A global training data set of 43340 pixels was thus created for testing the algorithm. Table 10 shows the number of training samples used for each surface type. The Water Bodies could not be tested because the AVHRR data contain a land/water mask.

Table 10. Number of training pixels used in development and testing of the VIIRS prototype 1 km product.

IGBP Class Name	Number of Training Samples
1) Evergreen Needleleaf Forests	1806
2) Evergreen Broadleaf Forests	4658
3) Deciduous Needleleaf Forests	504
4) Deciduous Broadleaf Forest	2091
5) Mixed Forests	1855
6) Closed Shrbulands	1412
7) Open Shrublands	3246
8) Woody Savannas	4247
9) Savannas	1907
10) Grasslands	4211
11) Permanent Wetlands	630
12) Croplands	6907
13) Urban/Built-Up	252
14) Croplands/Natural Vegetation Mosaics	4000
15) Snow and Ice	1705
16) Barren	3909
17) Water Bodies	Land/Water Mask

The training data listed above are generally evenly distributed over the surface of the Earth and are available for every continent except Antarctica. Because no *a priori* reliable estimates of the actual distribution of IGBP cover types on the Earth is available, the numbers in Table 10 may not exactly reflect this actual distribution.

To generate the algorithm error specification, the typing accuracy for the VIIRS Quarterly Surface Types IP was tested on this global sample of 43,340 training points known *a priori* to represent 16 of the 17 IGBP surface types. The full 43,340 global training samples were randomly split into an 80% training sample which was used to develop a boosted decision tree which was then applied to the remaining 20% (8639 samples) of the samples for testing, following Strahler *et al*. (1999). The overall mean typing accuracies, and their standard deviations, obtained when applying the boosted decision trees to the testing samples for 50 such random training/testing replications were retained and represent a robust approximation of the true accuracy, and its potential true variability. An overall mean typing accuracy of 88% (±0.25% standard deviation) was obtained from the 50 replicated sets. The summary stratification of the mean correct typing probability for these same tests over the 16 classes considered is as follows:

- Evergreen Needleleaf Forests: 76.04% (±2%)
- Evergreen Broadleaf Forests: 95.21% (±0.5%)

- Deciduous Needleleaf Forests: 85.15% (±3.6%)
- Deciduous Broadleaf Forests: 79.94% (±1.7%)
- Mixed Forests: 76.62% (±1.9%)
- Closed Shrublands: 81.63% (±2.2%)
- Open Shrublands: 92.46% (±1%)
- Woody Savannas: 79.69% (±1.1%)
- Savannas: 59.63% (±2.6%)
- Grasslands: 88% (±0.9%)
- Permanent Wetlands: 74.19% (±3.7%)
- Cropland: 93.58% (±0.5%)
- Urban/Built-up: 43.76% (±6.7%)
- Cropland/Natural Vegetation Mosaics: 94.42% (±0.8%)
- Snow/Ice: 99.64% (±0.4%)
- Barren: 98.96% (±0.4%)
- Water Bodies: Mask used

The overall typing accuracy of 88% is well above the threshold requirement of 70%. In addition, a standard deviation of a quarter of a percent implies that these accuracies are very robust and stable. Threshold performances are achieved in 14 of the 16 classes considered, with only the Urban class, as expected, and the Savannas class having accuracies lower than 70%. Further examination and refinement of the Savannas class will be performed in order to determine the reasons for this lower accuracy. The VIIRS low light level sensor may be required to produce a city lights product that can be used to better determine Urban areas of the Earth. In general, the results obtained for all of the other cover types indicate that the boosted decision tree is reliable, stable, and accurate.

Figure 15 shows the prototype VIIRS Surface Type global product that is produced when the boosted decision tree is applied to the remaining AVHRR data. This is the first such product to be produced in an entirely automated fashion worldwide. The accuracies listed above are also the best that have been retrieved to date on similar global products. These advances made on AVHRR products could only been achieved by the use of the global training data set published by the University of Maryland and the judicious development of temporal metrics which amplify the separation of coarse scale cover types. The C5.0 decision tree algorithm is the tool that allows us to make this a flexible, robust, accurate, and automated algorithm that is tailored for operational use.

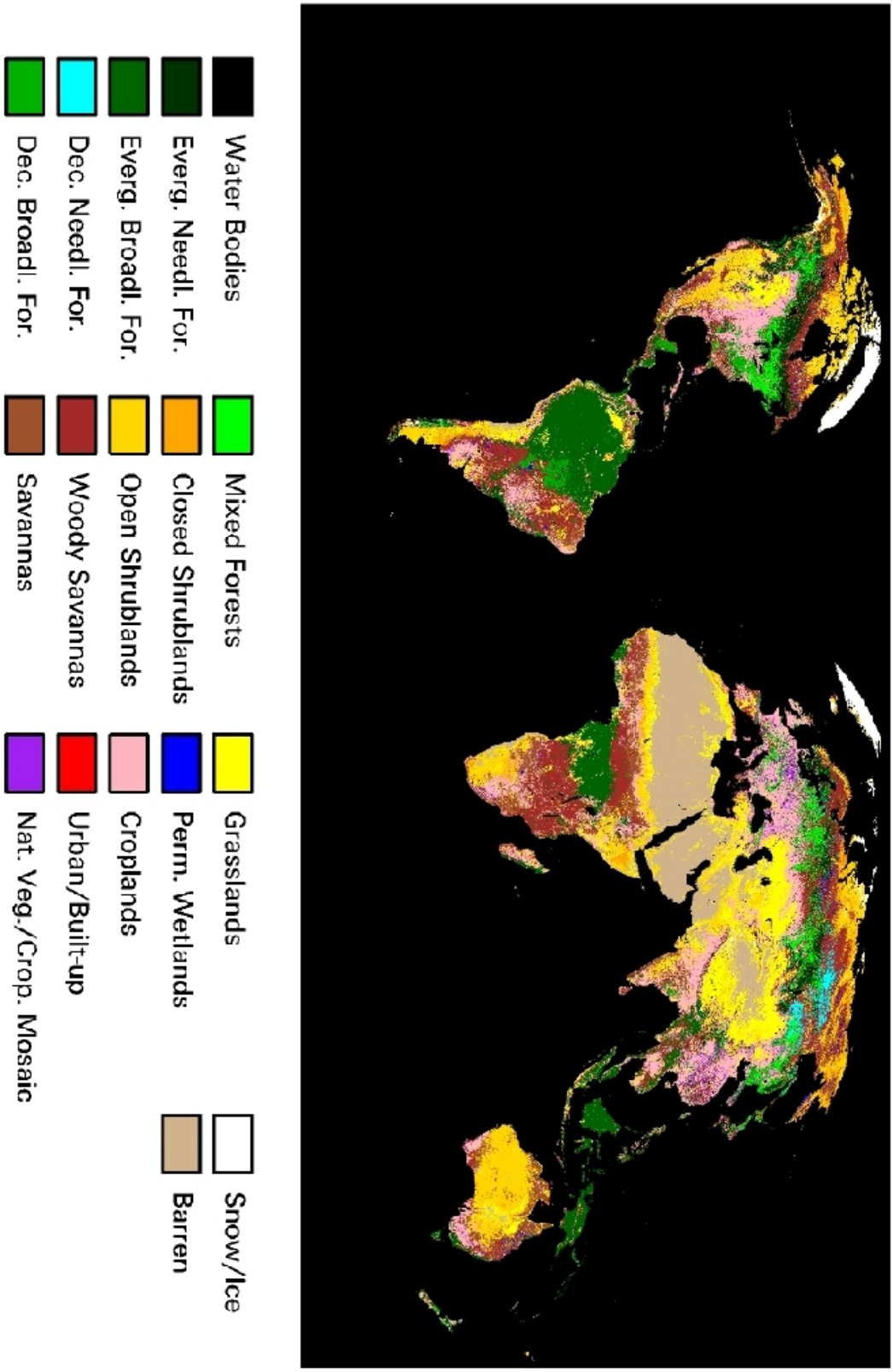


Figure 15. Prototype VIIRS Surface Type Global product.

6.0 REFERENCES

Abuelgasim, A.A., S. Gopal, J.R. Irons, and A.H. Strahler (1996). Classification of ASAS multiangle and multispectral measurements using artificial neural networks. *Remote Sens. Environ.*, 57:79-87.

Ackerman, S. *et al.* (1997). Discriminating clear-sky from cloud with MODIS: Algorithm Theoretical Basis Document, Version 3.2., Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Asrar, G., M. Fuchs, E. T. Kanemasu, and J. L. Hatfield (1984). Estimating absorbed photosynthetically active radiation and leaf area index from spectral reflectance in wheat. *Agron. J.*, 76:300-306.

Bonan, G.B. (1995). Land-atmosphere interactions for climate system models: Coupling biophysical, biogeochemical, and ecosystem dynamical processes. *Remote Sens. Environ.*, 51:57-73.

Charney, J., P.H. Stone, and W.J. Quirk (1975). Drought in the Sahara: a biogeophysical feedback mechanism. *Science,* 187:434-435.

Condit, H. R. (1970), The spectral reflectance of american soils, *Photogramm. Eng. Rem. Sens*., 36:955-966.

Daughtry, C. S. T., K. P. Gallo, S. N. Goward, S. D. Prince, and W. P. Kustas (1992). Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies, *Remote Sens. Environ.*, 39:141-152.

DeFries, R. S., and J. R. G. Townshend (1994a). Global land cover: comparison of ground-based data sets to classifications with AVHRR data. In Foody, G. and P. Curran (eds.*) Environmental Remote Sensing from Regional to Global Scales*, pp. 84-110, John Wiley and Sons, Chichester.

DeFries, R. S., and J. R. G. Townshend (1994b). NDVI-derived land cover classifications at a global scale. *Int. J. Rem. Sens.*, 15:3567-3586.

DeFries, R. S., M. Hansen, and J. Townshend (1995a). Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Rem. Sens. Environ.*, 54:209-222.

DeFries, R. S., *et al.* (1995b). Mapping the land surface for global atmosphere-biosphere models: Towards continuous distributions of vegetation's functional properties. *J. Geophys. Res.*, 100:20867-20882.

DeFries, R., M. Hansen, M. Steininger, R. Dubayah, R. Sohlberg, and J. Townshend (1997). Subpixel forest cover in central Africa from multisensor, multitemporal data. *Rem. Sens. Environ.*, 60:228-246.

DeFries, R. S., M. Hansen, J. R.G. Townshend, and R. Sohlberg (1998a). Global land cover classifications at 8 km spatial resolution: the use of training data derived from landsat imagery in decision tree classifiers. *Int. J. Rem. Sens.*, 19:3141-3168.

DeFries, R. S., J. R.G. Townshend, and M. Hansen (1998b). Continuous fields of vegetation characteristics at the global scale. *Int. J. Rem. Sens.*, in press.

Dickinson, R.E. (1983). Land surface processes and climate-surface albedos and energy balance. *Advances in Geophysics*, 25:305-350.

Dickinson, R.E. (1995). Land processes in climate models. *Rem. Sens. Environ*, 51:27-38.

Dickinson, R. E., A. Henderson-Sellers, P.J. Kennedy, and M. F. Wilson (1986). *Biosphere-atmosphere transfer scheme (BATS) for the NCAR Community Climate Model*, NCAR Tech Note NCAR/TN-275+STR.

Ehrlich, D., and E.F. Lambin (1996). Broad scale land-cover classification and interannual climatic variability. *Int. J. Rem. Sens.*, 17:845-862.

Food and Agriculture Organization of the United Nations, *Digital Soil Map of the World and Derived Soil Properties-Version 3.5*, FAO/UNESCO, Rome, 1995.

Friedl, M. A., and C.E. Brodley (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.*, 61:399-409.

Graetz, D. (1994). Grasslands. In W. B. Meyer, and B. L. Turner III. (eds.) *Changes in land use and land cover: A global perspective*, New York, Cambridge University Press, NY. pp. 125-148.

Gumuzzio, J., et al. (1997), Study to investigate land degradation applications using a high resolution land surface processes mission, ESA Study Contract Report, REP/09/97/1.

Gutman, G. and A. Ignatov (1998), The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models, *Int. J. Remote Sens.*, 19:1533-1543.

Hall, D. K. et al., 1999, Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow, Lake Ice- and Sea Ice-Mapping Algorithms, *Algorithm Theoretical Basis Document, Version 4.0.,* Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Hansen, M., R. Dubayah, and R. DeFries (1996). Classification trees: an alternative to traditional land cover classifiers. *Int. J. Remote Sens.*, 17:1075-1081.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg (1998). Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Rem. Sens.* 21: 1331-1364.

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Rem. Sens.*, 12:1147-1163.

Holben, B.N., Y. Kaufman and J. Kendall (1990). NOAA-11 AVHRR visible and near-IR inflight calibration. *Int. J. Rem. Sens.*, 11:1511-1519.

Houghton, R. A., and G. M. Woodwell (1989). Global climatic change, *Scientific American*, 260:36-44.

Huete, A. R., 1988, A soil adjusted vegetation index (SAVI). *Remote Sens. Environ.*, 25: 295-309.

Irons, J. R., R. A. Weismiller, and G. W. Petersen (1989), Soil Reflectance, in Asrar, G. (ed.*), Theory and Applications of Optical Remote Sensing*, pp. 66-106, John Wiley and Sons, New York.

Jacquemoud, S., F. Baret, and J. F. Hanocq (1992), Modeling spectral and bidirectional soil reflectance, *Rem. Sens. Environ.,* 41:123-132.

James, M. E., and S. N. V. Kalluri (1994). The pathfinder AVHRR land data set: an improved coarse resolution data set for terrestrial monitoring, *Int. J. Rem. Sens.*, 15:3347-3364.

Justice, C. O. *et al.* (1998) The Moderate Resolution Imaging Spectroradiometer (MODIS): Land Remote Sensing for Global Change Research. *IEEE Trans. Geosci. Rem. Sens*., 36:1228-1249.

Justice, C. O., J. R. G. Townshend, B. N. Holben, and C. J. Tucker (1985). Analysis of the phenology of global vegetation using meteorological satellite data. *Int. J. Rem. Sens.*, 6:1271-1318.

Kealy, P. S. and S. J. Hook (1993), Separating temperature and emissivity in thermal infrared multispectral scanner data: implications for recovering land surface temperature, *IEEE Trans. Geosci. Rem. Sens.,* 31:1155-1164.

Kohonen, T. (1997). *Self-Organizing Maps*, Springer-Verlag, Berlin.

Lambin, E.F., and D. Ehrlich (1996). The surface-temperature-vegetation index space for land cover and land-cover change analysis. *Int. J. Rem. Sens.*, 17:463-487.

Laporte, N., C. Justice, and J. Kendall (1995). Mapping the dense humid forest of Cameroon and Zaire using AVHRR data. *Int. J. Rem. Sens.*, 16:1127-1145.

Laporte, N. T., S.J. Goetz, M. Heinicke, and C.O. Justice (1998). A new land cover map of central africa derived from multi-resolution, multi-temporal AVHRR data. *Int. J. Rem. Sens.*, in press.

Lloyd, D. (1990). A phenological classification of terrestrial vegetation using shortwave vegetation index imagery. *Int. J. Rem. Sens.*, 11:2269-2279.

Los, S. O, C. O. Justice, and C. J. Tucker (1994). A global 1° by 1° NDVI data set for climate studies derived from the GIMMS continental NDVI data. *Int. J. Rem. Sens.*, 15:3493-3518.

Loveland, T. R., J. W. Merchant, D. O. Ohlen and J. F. Brown (1991). Development of a land-cover characteristics database for the conterminous U.S. *Photogramm. Eng. Remote Sens.*, 57:1453-1463.

Loveland, T.R., and A.S. Belward (1997). The IGBP-DIS global 1 km land cover data set, DISCover: first results. *Int. J. Rem. Sens.*, 18:3289-3295.

Loveland, T.R., *et al.* (1999), An analysis of the IGBP global land-cover characterization process, *Photogramm. Eng. Remote Sens.*, 65:1021-1032.

Matthews, E., 1983, Global Vegetation and Land Use; New High-Resolution Data Bases for Climate Studies, *Journal of Climate and Applied Meteorology,* 22:474-487.

Melillo, J. M. (1994). Modeling land-atmosphere interactions: A short review. In W. B. Meyer and B. L. Turner III (eds.) *Changes in land use and land cover: A global perspective* New York; Cambridge University Press, pp. 387-410.

Melillo, J.M. *et al.* (1993). Global climate change and terrestrial net primary production. *Nature*, 234:-240.

Muchoney, D., *et al.* (2000). Application of the MODIS global supervised classification model to vegetation and land cover mapping of Central America. *Int. J. Rem. Sens.*, 21:1115-1138.

Mueller-Dombois, D. (1984). Classification and mapping of plant communities: a review with emphasis on tropical vegetation. In G.M. Woodwell (ed.) *The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing*, John Wiley and Sons, Chichester.

Nemani, R., and S.W. Running (1995). Satellite monitoring of global land cover changes and their impact on climate. *Climatic Change*, 31:395-413.

Nemani, R., and S.W. Running (1996). Implementation of a hierarchical global vegetation classification in ecosystem function models. *J. Veg. Sci.*, 7:337-346.

Nemani, R., and S.W. Running (1997). Land cover characterization using multitemporal red, near-Ir, and thermal-Ir data from NOAA/AVHRR. *Ecological Applications*, 7:79-90.

Olson, J. S., Watts, J., and L. Allison (1983). Carbon in live vegetation of major world ecosystems. W-7405-ENG-26, U.S. Department of Energy, Oak Ridge National Laboratory.

Prentice, K.C. (1990). Bioclimatic distribution of vegetation for general circulation models. *J. Geophys. Res.*, 95:11811-11830.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*, Morgan Kaufman Publishers, Inc., San Mateo, CA.

Richards, J. A. (1983). Remote sensing digital image analysis. Springer-Verlag, Berlin.

Rivard, B., R. E. Arvidson, I. J. Duncan, M. Sultan, and B. El. Kaliouby(1992), Varnish, sediment and rock controls on spectral reflectance of outcrops in arid regions, *Geology*, 20:295-298.

Running, S. W., and E. R. Hunt Jr. (1993). Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models, in Ehleringer, J.R. and C. Field, *Scaling Processes between leaf and landscape levels*, Academic Press, San Diego, pp. 141-158.

Running, S. W. *et al.* (1994). Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *Int. J. Rem. Sens.*, 15:3587-3620.

Running, S. W., T. R. Loveland, L. L. Pierce, R. R. Nemani and E. R. Hunt Jr. (1995). A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sens. Environ.*, 51: 39-48.

Running, S. W., R. Nemani and J.M. Glassy (1996). MODIS PSN (Net Photosynthesis) and MODIS NPP (Net Primary Productivity) products: *Algorithm Theoretical Basis Document, Version 3.0.*, Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Salisbury, J. W., and D. D'Aria (1992a). Emissivity of terrestrial materials in the 8-14 μm atmospheric window. *Rem. Sens. Environ.,* 42:83-106.

Salisbury, J. W., and D. D'Aria (1992b). Infrared (8-14 μm) remote sensing of soil particle size. *Rem. Sens. Environ.,* 42:157-165.

Scepan, J. (1999), Thematic validation of high-resolution global land-cover data sets. *Photogramm. Eng. Remote Sens.*, 65:1051-1060.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall (1994). A global 1° by 1° NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *Int. J. Rem. Sens.*, 15:3519-3546.

Shukla, J., C. Nobre, and P. Sellers (1990). Amazon deforestation and climate change. *Science*, 247:1322-1325.

Skole, D. L., and C. J. Tucker (1993). Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science*, 260:1905-1910.

Soil Survey Staff (1975), *Soil Taxonomy*, U.S. Government Printing Office, Washington, D.C.

Stoner, E. R. and M. F. Baumgardner, Characteristic variations in reflectance of surface soils, *Soil Sci. Soc. Am. J.,* 45:1161-1165, 1981.

Strahler, A.H. *et al.* (1996a). *MODIS Land Cover Product Algorithm Theoretical Basis Document, Version 4.1.*, Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Strahler, A. H., M. J. Barnsley, R. D'Entremont, B. Hu, P. Lewis, X. Li, J. P. Muller, C. B. Schaaf, W. Wanner, and B. Zhang (1996b). *MODIS BRDF Albedo Product: Algorithm Theoretical Basis Document, Version 3.2.*, Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Townshend, J. R. G. (1999). MODIS Enhanced Land Cover and Land Cover Change Product *Algorithm Theoretical Basis Document, Version 2.0.*, Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd/modistables.html.

Townshend, J. R. G., C. O. Justice, and V. T. Kalb (1987). Characterization and classification of South American land cover types using satellite data. *Int. J. Remote Sens.*, 8:1189-1207.

Townshend, J. R. G., C. O. Justice, C. Gurney, and J. McManus (1992). The impact of misregistration on change detection. *IEEE Trans. Geosci. Rem. Sens*., 30:1054-1060.

Townshend, J. R. G., C. O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski, and S. Ruttenberg (1994). The 1 km resolution global data set: needs of the International Geosphere Biosphere Programme. *Int. J. Rem. Sens.*, 15:3417-3442.

Townshend, J. R. G., C. O. Justice, W. Li, C. Gurney and J. McManus, 1991, Global land cover classification by remote sensing: present capabilities and future possibilities, *Remote Sens. Environ.*, 35:243-255.

Tucker, C. J., J. R. G. Townshend, and T. E. Goff (1985). African land-cover classification using satellite data. *Science*, 227:369-375.

Vermote, E. (1996). *Atmospheric Correction Algorithm: Spectral Reflectances*: *Algorithm Theoretical Basis Document, Version 3.0.*, Available from World Wide Web Site: http://spso.gsfc.nasa.gov /atbd /modistables.html.

Walthall, C. L., and E. C. Brown de Colstoun (1997). Medium altitude bidirectional imaging spectroradiometer measurements over the HAPEX-Sahel west central super site using the advanced solid-state array spectroradiometer (ASAS). In P. Kabat, S. D. Prince, and L. Prihodko (eds.), *HAPEX Sahel west central supersite: Methods, measurements and selected results*, Report 130, DLO Winand Staring Centre, Wageningen, The Netherlands, in press.

Weiss, S.M. and C. A. Kulikowski (1991), *Computer systems that learn*, Morgan Kaufman Publishers, San Mateo.

Williams, M. (1994). Forest and tree cover. In W. B. Meyer and B. L. Turner III (eds.) *Changes in land use and land cover: A global perspective*. New York; Cambridge University Press, pp. 97-124.

Wilson, M.F., and A. Henderson-Sellers (1985). A global archive of land cover and soils data for use in general circulation models. *J. Clim*., 5:119-143.

Wu, A., Z. Li, and J. Cihlar (1995). Effects of land cover type and greenness on advanced very high resolution radiometer bidirectional reflectances: analysis and removal. *J. Geophys. Res.*, 100: 9179-9192.

APPENDIX A ANALYSIS OF THE DIRECT RETRIEVAL OF SOIL CLASSES FROM VIIRS OPTICAL/THERMAL DATA

A.1 Soil Class Retrieval - Introduction

As described in the old version of the SRD (p. 40), the "Surface Type is defined as the predominant vegetation and/or soil type in a given area". One of the issues regarding the IPO classification scheme is that six out of the 21 surface types appear to be some combination of soil texture types or at least a gradation of particle sizes. The soil types in question are: Rocky Fields, Gravel, Sandy Soils, Loam, Clay, and Peat. The purpose of having these soil classes as surface types is presumably to serve as inputs to Land Surface Parameterization schemes for climate modeling because of their effect on sensible heat fluxes, runoff, and soil water storage, but they could also serve some military tactical purpose.

At this time, the direct retrieval of soil type and/or texture from visible/infrared and thermal remotely sensed data has not been demonstrated at the regional and global scales, and much less in an operational scenario. Additionally, no current or planned land cover classification product includes soil types as land cover classes. After discussions with our science team, we had serious concerns that these soil surface types could not be retrieved even at the threshold level and as such we had proposed the usage of a static digital soil database available from the Food and Agriculture Organization (FAO) in Rome (FAO, 1995).

In this section we document a series of analyses performed to assess the feasibility of extracting soil type information directly from optical/thermal data acquired by VIIRS. Results using laboratory acquired soil spectra show that soil types and general groupings can have a rather large natural variability in reflectances at many wavelengths in the spectrum which does not appear strongly correlated to soil texture components. In addition, no reliable trends were found to support the accurate separation of even fairly broad soil categories. However, analyses using a decision tree algorithm did show some moderate success in separating sandy, loamy, and clayey soils. This type of analysis may warrant further exploration in the future with expanded soil types and/or simulated satellite data.

A.2 Soil Class Retrieval - Background

One interpretation of the "and/or" in the SRD sentence above is that the IPO want the vegetation AND associated soil types within a pixel or cell. This could mean that a Tropical Forest could be on clayey soils or a Desert on sandy soils, for example. Another interpretation is that only areas that are bare at any time of the year should be classified as the appropriate soil type. This would mean that annual crops grown on sandy soils could be classified as Croplands in the global product and as Sandy Soil in the regional product in winter.

Under densely or even sparsely vegetated conditions, the direct retrieval of background or soil information from the remotely-sensed data would be possibly achievable only through the use of an un-mixing algorithm, which is an objective at this time. Alternatively, if the soil information is to be provided through the usage of a static soil database, then data acquired by VIIRS will not be needed. Thus, there would appear to be some contradictions between the need for soil classes and the threshold and objective requirements.

In the case of bare soil surfaces, however, the combined optical/thermal properties of the soils may allow the direct separation of several broad soil types without having to depend on the soils database. We have submitted two Issue/Clarification/Support Request Forms (ICSR) in early March with questions addressing our interpretation of the Surface Type Environmental Data Record (EDR 40.6.4) requirements, as well as our need for definitions of the surface types.

One of the issues regarding the IPO soil types is that six out of the 21 surface types appear to be some combination of soil texture types, or at least a gradation of particle sizes, which makes their classification even more difficult. The soil types in question are: Rocky Fields, Gravel, Sandy Soils, Loam, Clay, and Peat. As shown in Table A.1, the U.S. Department of Agriculture (USDA) defines several classes of soil separates based on their particle size fractions. As we understand the IPO soil types, the Rocky Fields surface type would be composed of the coarsest materials, followed by Gravel, Sandy Soils, and so forth.

The USDA also uses a texture triangle to define soil textural classes based on the proportions of sand, silt, and clay present in the soil (Figure A.1). As can be seen, loams are typically made up of approximately equal amounts of sand, silt, and/or clay, and the corners of the triangle represent more or less pure sand, silt, or clay. Peat, although not a soil texture class *per se*, is typically characterized by very dark soils with a very high organic matter and water contents, and much smaller amounts of the other three materials.

Some key points can be made from Table A.1 and Figure A.1 regarding the IPO soil classes. If the classification is made according to the dominant material criteria, then many clayey soils will be classified as sandy. Likewise, a loamy soil will almost never be the dominant type because it is made up of the other three materials. Therefore, a percentage of either sand, silt, or clay will be greater than the other and, according to the IPO definitions, should be classified as such. Also, if the soil type definitions are based on soil texture then the silt category should be included instead of the Loam class. These issues may arise from the lack of definitions that we have been provided and we anticipate that with responses to our ICSRs they should be better resolved.

Intuitively, it would seem that sandy soils could be expected to be brighter in the visible/infrared portions of the spectrum than clayey soils, and in many cases this is indeed the case (Jacquemoud *et al*., 1992). However, soil reflectances result from a complex interplay between the reflective properties of constituents of the parent material, particle coatings, weathering processes, structure, texture, and moisture and organic matter contents (Irons *et al*., 1989). It is thus very difficult to verify a linkage between specific soil characteristics and the soil reflectance.

Table A.1. Soil particle fractions as defined by the U.S. Department of Agriculture (USDA) (Adapted from Irons *et al.* 1989).

Soil Separate	Particle Size Range (mm)
Gravel	> 200
Sand	0.05 - 1.00
Silt	0.002 - 0.05
Clay	< 0.002

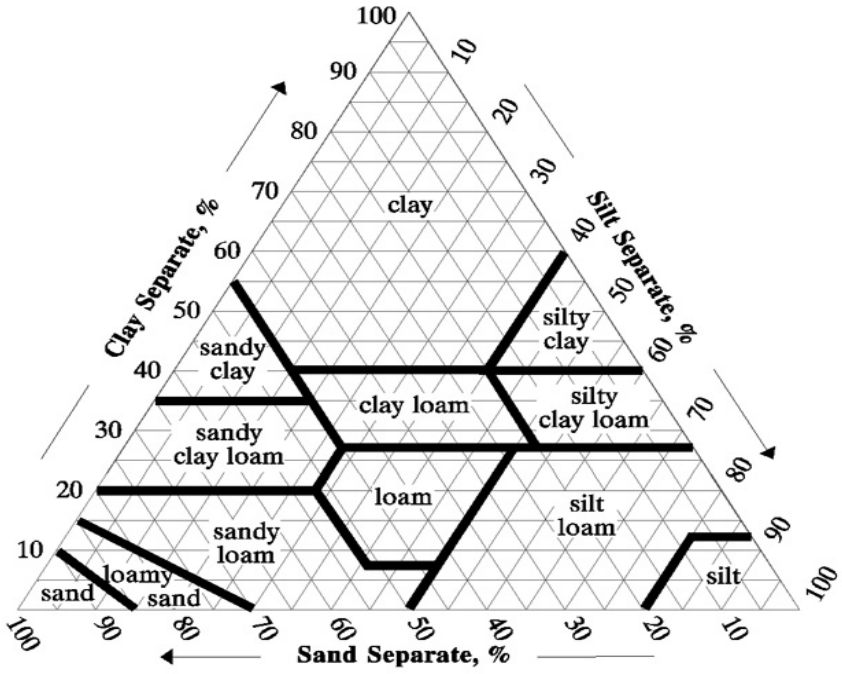


Figure A.1. USDA soil textural triangle. The percentage of sand, silt, and clay present in each soil textural class is obtained by reading the numbers parallel to each base of the triangle (i.e., percentages of clay are on lines parallel to the sand base).

Several attempts at providing general classifications based on soil spectral measurements have been made. Condit (1970) measured the spectra of 285 soils from the U.S. in the 320 to 1000 nm region and found that three general spectral curve shapes described the spectral curves from his database fairly accurately. Stoner and Baumgardner (1981) acquired spectral data from 485 soil samples from the U.S. and Brazil and recognized five basic soil reflectance curve shapes from 0.3 to 3.0 microns. The shapes were principally modified by the amount of organic matter and iron oxide present in the soils.

More typically, studies have noted weak or very weak relationships between soil properties such as sand or organic matter contents and reflectances in selected wavelength regions (Gumuzzio *et al*., 1997). Recently, the thermal properties of minerals, particularly quartz, and soils have been explored for the purpose of soil type discrimination (Salisbury and D'Aria, 1992a, 1992b). In the context of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) mission, Salisbury and D'Aria (1992b) showed that a ratio of ASTER bands 10 (8.125-8.475 μm) to 14 (10.95-11.65 μm) had a good, albeit non-linear, relationship to the combined amount of silt and clay present in the soil. However, soils with high organic matter or iron oxide contents were filtered through the use of another thermal band ratio between ASTER bands 10 and 12 (8.925-9.275 μm), and the authors expressed some concerns about the additional effects of soil moisture on these ratios. The general idea behind these studies is that the strong response of quartz, and by extension sand, in the 8 to 10 micron region could provide a basis for the discrimination of sandy from non-sandy soils. Because our nominal design for VIIRS contains a spectral band centered at about 8.55 μm, we felt that this type of approach could possibly be used to retrieve at least some soil information directly from the VIIRS data.

It cannot be stressed enough that the above studies were largely performed in the laboratory and under controlled conditions. Field and/or satellite remote sensing studies of soils will in all likelihood be affected by atmospheric, surface temperature, and bidirectional effects, among others, and most importantly, the natural variability of soil properties within the field of view of the sensor (Irons *et al*., 1989; Kealy and Hook, 1993). In fact, the retrieval of soil type and/or texture from visible/infrared and thermal remotely-sensed data has not been demonstrated at the regional and global scales, and much less in an operational scenario. We have very serious concerns that soil texture or soil classes cannot be retrieved at the threshold level and this is why we have proposed the usage of the FAO Digital Soil Map of the World (FAO 1995).

Nevertheless, we have taken an objective point of view on this issue and explored the information content of soil reflectances in an attempt to identify spectral regions whose combined information could yield a potential soil type retrieval algorithm. For this task, we acquired the spectra of 43 soil types from the United States and the Middle East directly from the ASTER spectral library. These spectra were measured by scientists at Johns Hopkins University (JHU), span the spectral range from 0.4 to 14 microns, and are representative of broad soil categories. The instrumentation and measurement techniques used to acquire the spectra in the laboratory are described in Salisbury and D'Aria (1992b) but can also be found at *http://asterweb.jpl.nasa.gov/speclib/jhu_description.html*. Table A.2 lists the principal attributes of the 43 soil types from the ASTER spectral library.

Table A.2 Description of soils from the Johns Hopkins University Spectral Library.The soil spectra are available through the ASTER internet site (*http://asterweb.jpl.nasa.gov/speclib/*).

NO.	NAME	CLASS	SUBCLASS	%Sand	%Silt	%Clay	%Organic
1	Reddish brown fine sandy loam	Alfisol	Paleustalf	85.5	7.7	7.3	0.16
2	Brown loamy fine sand	Alfisol	Haplustalf	85.2	8.5	6.3	0.86
3	Dark reddish brown fine sandy loam	Alfisol	Paleustalf	73.3	17.9	8.8	1.44
4	Brown sandy loam	Alfisol	Paleustalf	62.4	25.4	12.2	0.71
5	Brown to dark brown gravelly loam	Alfisol	Haploxeralf	60.2	36.5	4.2	8.54
6	Reddish brown fine sandy loam	Alfisol	Paleustalf	53.8	26.1	20.1	1.3
7	Brown fine sandy loam	Alfisol	Haplustalf	49.8	43.2	7	0.76
8	Brown fine sandy loam	Alfisol	Haplustalf	48.6	45	6.4	1.15
9	Pale brown silty loam	Alfisol	Fragiboralf	17.4	74.7	7.9	0
10	Very dark grayish brown loamy sand	Aridisol	Torripsamment	88.5	6.9	4.6	0.51
11	Light yellowish brown loamy sand	Aridisol	Camborthid	87.7	8.3	4	0.13
12	Brown gravelly sandy loam	Aridisol	Haplargid	81.3	15.5	3.2	0.1
13	Dark brown sandy loam	Aridisol	Calciorthid	79.8	15.5	4.7	0.48
14	Light yellowish brown interior dry gravelly loam	Aridisol	Calciorthid	48.9	40.4	10.7	0.5
15	Very pale brown to brownish yellow interior dry gravelly silt loam	Aridisol	Gypsiorthid	41.9	41.7	16.4	0.47
16	Dark brown interior moist clay loam	Aridisol	Salorthid	38.4	50.5	11.1	1.03
17	Light yellowish brown clay	Aridisol	Salorthid	29.7	18.8	51.5	0.68
18	Brown silty loam	Aridisol	Camborthid	28.5	64.3	7.2	0.89
19	Dark yellowish brown silty clay	Aridisol	Salorthid	26.6	15.3	58.1	0.4
20	Light yellowish brown loam	Aridisol	Calciorthid	11.2	70.7	18.1	1.02
21	White gypsum dune sand	Entisol	Torripsamment	100	0	0	0
22	Brown to dark brown sand	Entisol	Quartzipsamment	94	3.8	2.2	0.52
23	Brown to dark brown silt loam	Entisol	Ustifluvent	21.5	46.3	32.2	0.43
24	Brown sandy loam	Inceptisol	Haplumbrept	75.9	14.2	9.9	4.85
25	Dark brown fine sandy loam	Inceptisol	Haplumbrept	75.3	18.8	5.9	3.41
26	Gray/dark brown extremely stoney coarse sandy loam	Inceptisol	Cryumbrept	71.8	21.1	7.1	6.18
27	Brown to dark brown gravelly fine sandy loam	Inceptisol	Xerumbrept	68.3	29	2.7	5.43
28	Dark yellowish brown micaceous loam	Inceptisol	Dystrochrept	43.6	30.7	25.7	0.96
29	Pale brown dry silty clay loam	Inceptisol	Ustocrept	16.1	47.4	36.5	0.95
30	Very dark grayish brown silty loam	Inceptisol	Plaggept	6.1	84.1	12.9	3.8
31	Brown to dark brown sandy loam	Mollisol	Hapludoll	73.2	18.7	8.1	0.61
32	Grayish brown loam	Mollisol	Haplustall	61.3	26.7	12	2.49
33	Very dark grayish brown loam	Mollisol	Agriudoll	53.7	32.4	13.9	1.06
34	Very dark grayish brown loam	Mollisol	Cryoboroll	50.6	31.9	17.5	2.87
35	Very dark grayish brown loam	Mollisol	Paleustoll	44.5	41	14.5	1.11
36	Black loam	Mollisol	Cryoboroll	36.6	45.9	17.5	6.64
37	Very dark grayish brown silty loam	Mollisol	Argiustoll	26.9	50	23.1	1.57
38	Very dark grayish brown silty loam	Mollisol	Argiustoll	22.6	54.6	22.8	2.22
39	Gray silty clay	Mollisol	Haplaquoll	11.3	63.1	25.6	1.88
40	Dark grayish brown silty loam	Mollisol	Agialboll	6.3	63.4	30.3	2.29
41	Dark reddish brown, organic-rich, silty loam	Spodosol	Cryohumod	0	99.96	0.04	28.5
42	Brown to dark brown loamy sand	Ultisol	Hapludult	90.2	7.4	2.4	0.37
43	Brown to dark brown clay	Vertisol	Chromoxerert	34.4	27.4	38.2	1.4

A.3 Soil Class Retrieval - Results

A.3.1 General Soil Reflectance Features

Figures A.2 and A.3 show the full spectra of five "brown to dark brown" colored soils from the ASTER spectral library. These soils represent four soil classes and have varied quartz and sand contents (see Table A.2). Most of these soils have a fairly low organic matter content except for the gravelly loam soil which has an organic matter content of 8.54 percent. In the 0.4 to 3.0 micron spectral range (Figure A.2), there is a fairly large variation in soil brightnesses at almost all wavelengths, with a near-equal separation of the soil reflectances, apparently unrelated to sand content. The clay soil has similar reflectances to the sand in the 1.0 to 2.5 micron range while the sandy loam is much brighter than the sand for the entire visible/infrared range. Also, the soil with the highest organic matter content (gravelly loam) does not have a lower reflectance than the silt loam, as could be expected. The shape of this curve between 0.4 and 1.5 microns is, however, less concave than the shapes for the other soils, and is qualitatively similar to those organic-dominated soils described by Stoner and Baumgardner (1981).

From 3.0 to 14.0 microns, several prominent spectral regions of variability are evident (Figure A.3), such as those between 3.0 and 6.0 microns, and 8.0 and 10.0 microns, the latter being associated with the quartz restsrahlen bands (Salisbury and D'Aria, 1992a, 1992b). Additionally, a smaller region of interest is found between 12.5 and 13.5 microns and again is related to the amount of quartz present in the soil. The variability in reflectances found between 3.0 to 6.0 microns does not appear to follow any trends with sand content, with the silty loam, sandy loam, and clay all having greater reflectances than the sand. In the 8.0 to 10.0 micron region, decreasing reflectances with decreasing sand contents are apparent for four of the soils except the clay. In addition, the spectra for the gravelly loam and silt loam show the effects of organic matter and iron oxide on the two quartz restsrahlen peaks. According to Salisbury and D'Aria (1992b), organic matter tends to subdue the quartz peaks while iron oxide may distort the far peak as seen for the silt loam spectrum. This forms the basis for an ASTER band 10/12 ratio to separate those soils.

A.3.2 Soil Reflectance Variability in Nominal VIIRS Spectral Bands

We have also examined the variability or information content of the reflectances throughout the spectrum. For this task, we have averaged the hyperspectral ASTER reflectances to approximately conform to the nominal spectral bands of VIIRS as well as several of the ASTER or MODIS thermal bands, providing a nearly continuous spectral coverage of the soil reflectances. Table A.3 shows the bandwidths selected for this analysis. We have also stratified the ASTER soils into Sandy, Loamy, and Clayey soil categories according to the general groupings of soil textural classes given by the USDA (Soil Survey Staff, 1975):

- *Sandy soils:* Sands, Loamy Sands.
- *Loamy Soils:* Sandy Loam, Loam, Silt Loam, Silt, Clay Loam, Sandy Clay Loam, Silty Clay Loam.
- *Clayey Soils:* Sandy Clay, Silty Clay, Clay.

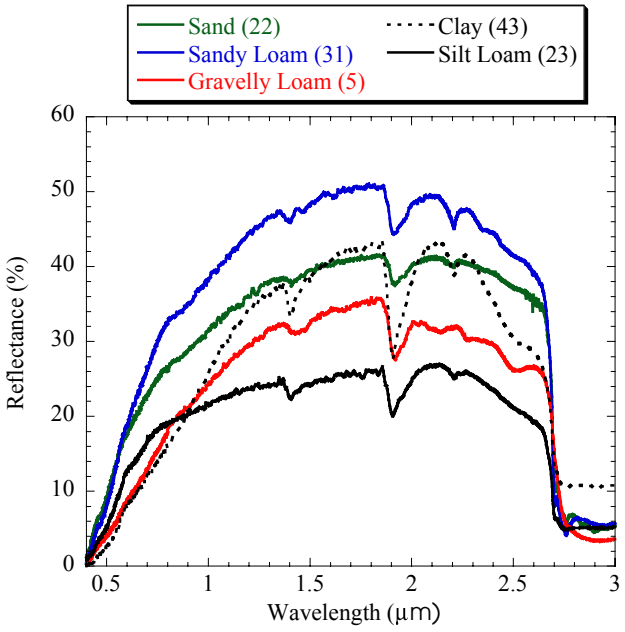


Figure A.2. Selected spectra from 0.4 to 3.0 µm for five brown to dark brown colored soils. Numbers in parentheses represent sample numbers given in Table A.2.

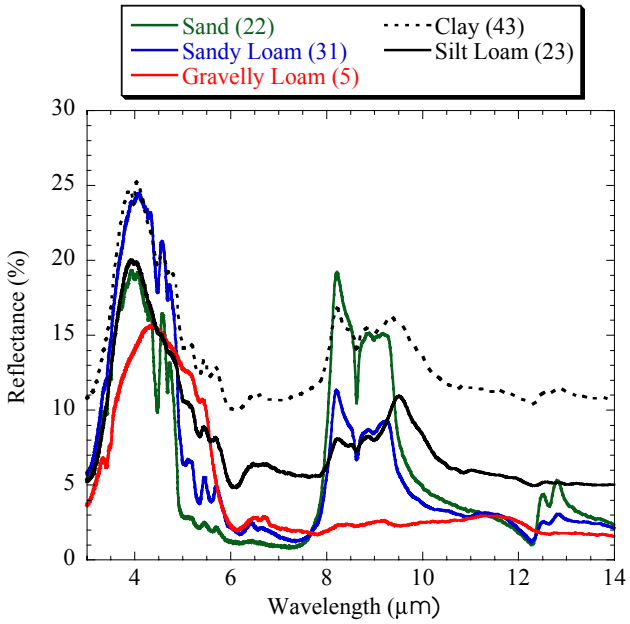


Figure A.3.Same spectra as Figure A.2 but covering the spectral range from 3.0 to 14.0 µm.

The USDA also provides quantitative definitions for each of the above soil types according to the percent sand, silt, and/or clay content of any soil. We have taken the soil types given in the ASTER library and grouped them into the three general groupings. From this grouping, 6 of the soils are classified as Sandy, 33 as Loamy, and the remaining 4 as Clayey. Unfortunately, the sample numbers are not equal, with the Loamy soils being quite dominant.

Figure A.4 shows the means and standard deviations for all soils, as well as the three general groupings described above. For this analysis the spectrum for the gypsum sand was omitted because it was an obvious outlier at practically all wavelengths. Unexpectedly, the means for the clayey soils are higher in the visible region than those for the sandy soils but this could be attributable to the small number of clay soil samples. Other studies have found that clayey soils tend to be generally darker in this wavelength region (Jacquemoud *et al*., 1992).

The variability between the mean reflectances for the different soil types appears to be similar for most of the visible and middle infrared regions and is relatively small, with 5-10 percent reflectance differences throughout the range. The standard deviations for the groupings are in most cases larger than these mean differences, although the standard deviations for the clayey soils are substantially higher when compared to the other soils, particularly the sandy soils. As seen for the individual spectra, there are also regions of reflectance variability at around 3.75 μm, and again from 8 to 10 microns. The decreasing trend of reflectances at 3.75 μm with decreasing sand content may hold some additional information to that provided by the 8-10 μm region. Between 8 to 10 microns the effect of the quartz restsrahlen bands is apparent for the sandy soils, and not prominent for the clayey soils, indicating that a thermal band ratio approach, as suggested by Salisbury and D'Aria (1992b), appears warranted in this case, and would help to separate these soil types.

Figure A.5 shows the coefficients of variation (CV), or the standard deviation normalized by the sample mean, for the same data as seen in Figure A.4. This was done to minimize the apparent dependence of the standard deviations on the sample reflectances. Again, these plots are suggestive of a fairly large information content, or at least a large reflectance variability, in several spectral regions, particularly in the visible and thermal. In the visible wavelengths, CVs are typically quite large, meaning that the deviations are equal to at least half of the sample means. This apparent natural variability also highlights the difficulties in ascribing a change in reflectance to any particular causative factor. For example, the variability in reflectances for the clay soils, as shown by the CV, is larger than the variability shown for all soils together. This is in part due to the small number of samples but yet it is much smaller for the Sandy Soils. The four clay soils all have fairly similar sand, silt, clay, and organic matter contents yet their reflectances in the visible can deviate by 15 to 30 percent (Figure A.6).

Table A.3. Bandwidths for spectral bands selected for the analysis. Some bands approximately correspond to the VIIRS nominal bands, others to ASTER and MODIS.

Band No.	Bandwidth (μm)	Approx. Band Center (μm)	Band Designation		
			VIIRS	MODIS	ASTER
1	0.402-0.422	0.412	VIIRS 1	MODIS 8	
2	0.433-0.453	0.443	VIIRS 2	MODIS 9	
3	0.453-0.480	0.466		MODIS 3	
4	0.480-0.500	0.490	VIIRS 3	MODIS 10	
5	0.500-0.545	0.523		MODIS 11	
6	0.545-0.565	0.555	VIIRS 4	MODIS 4	
7	0.565-0.620	0.593			
8	0.620-0.670	0.645	VIIRS 5	MODIS 1	
9	0.670-0.841	0.742			
10	0.841-0.876	0.860	VIIRS 6	MODIS 2	
11	0.876-0.931	0.904		MODIS 17	
12	0.915-0.965	0.940		MODIS 19	
13	0.965-1.230	1.098			
14	1.230-1.250	1.244		MODIS 5	
15	1.250-1.360	1.308			
16	1.360-1.390	1.376	VIIRS 7	MODIS 26	
16	1.390-1.628	1.376			
17	1.628-1.652	1.510	VIIRS 8	MODIS 6	
18	1.652-2.105	1.640			
19	2.105-2.155	2.130	VIRS 9	MODIS 7	
20	2.155-3.660	2.778			
21	3.660-3.840	3.749	VIIRS 10	MODIS 20	
22	3.840-3.929	3.885			
23	3.929-3.989	3.961		MODIS 21-22	
24	3.989-4.500	4.237			
25	4.500-5.000	4.743			
26	5.000-5.500	5.268			
27	5.500-6.000	5.769			
28	6.000-6.535	6.257			
29	6.535-6.895	6.713		MODIS 27	
30	6.895-7.500	7.195			
31	7.500-8.125	7.806			
32	8.125-8.475	8.296			ASTER10
33	8.400-8.700	8.549		MODIS 29	
34	8.700-8.925	8.817			
35	8.925-9.275	9.072			ASTER12
36	9.275-10.250	9.717			
37	10.250-10.750	10.496	VIIRS 11		ASTER13
38	10.780-11.280	11.032	VIIRS 11	MODIS 31	ASTER14
39	11.280-11.770	11.535			
40	11.770-12.270	12.030	VIIRS 12	MODIS 32	
41	12.270-13.185	12.727			
42	13.185-13.485	13.345			
43	13.485-13.785	13.661			
44	13.785-14.000	13.889			

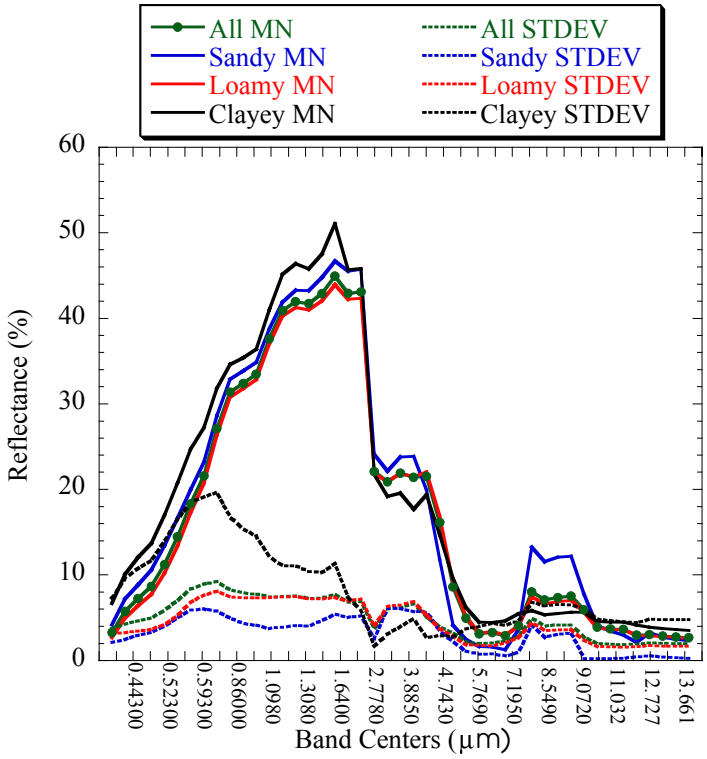


Figure A.4. Means (MN) and standard deviations (STDEV) in reflectances for bands in Table A.3 for all soils and for three general soil groupings. The x-axis values have been compressed for clarity and only represent the approximate positions of band centers.

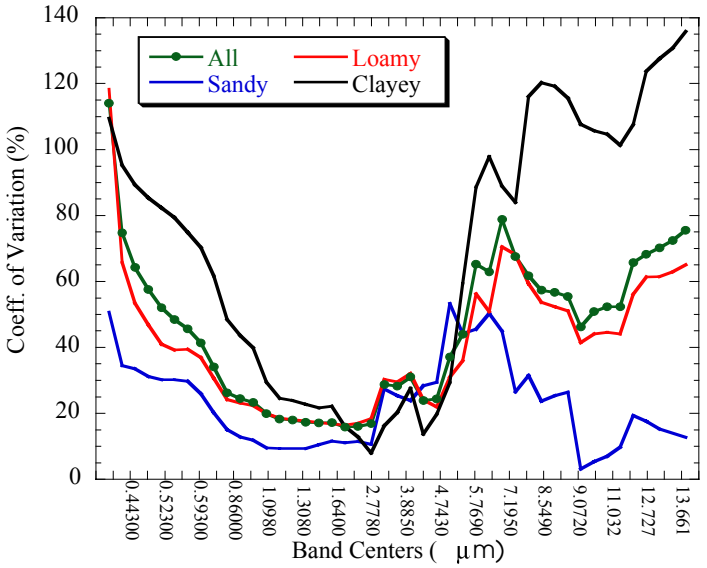


Figure A.5. Coefficients of Variation for same data presented in Figure A.4.

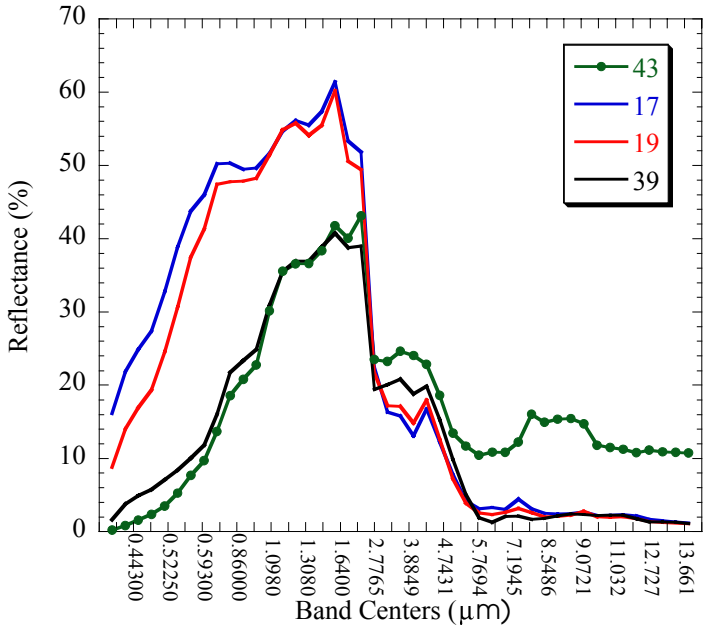


Figure A.6. Mean reflectances for the four clayey soils. Numbers in legend correspond to sample numbers in Table A.2.

For all soils, CV values greater than 50 percent are also found at 7.2 µm and throughout the thermal regions, further highlighting the potential soil information in this spectral region. For the sandy soils, less variation is seen in the 8-10 micron range, presumably because of the high sand content for all these soils, while the loamy soils show larger variability in this region because they encompass a greater variety of soil types. Again, large CV values are found in the thermal wavelengths for the clayey soils. It should be noted that similar observations were made when the soils were stratified by soil class (i.e., Alfisol, Aridisol, Inceptisol, Mollisol), although in this case the Inceptisols had CVs greater than 100 percent at almost all wavelengths. These observations suggest that in many cases the within group variations may be greater than the variation between groups and raise some important concerns about the potential retrieval of soil information from remotely-sensed data. However, if general relationships can be developed that explain a large amount of the variations, and if it can be determined what the variability is related to, a potential may still exist to separate these broad soil categories.

A.3.3 Linear Regression Analysis

We have explored the relationships between the spectral reflectances at all bands to the different soil components (i.e., sand, silt, clay, and organic matter contents) in order to develop potential metrics for soil type separation. The general idea was that if some general thresholds could be found, they could be used to discriminate between the three broad soil types.

Extremely weak or no linear relationships were found between sand, silt, and clay contents and reflectances at all bands in Table A.3. Some moderate correlation (r^2>0.3), however, was found between sand and silt contents at several bands between 4.75 and 6.0 microns. This is unfortunately an area of very strong atmospheric gaseous absorption due to CO_2. For the bands between 8 and 10 microns, r^2 values were between 0.2 and 0.3 with maximum values at 8.55 and

8.8 µm (r^2=0.29 and 0.3, respectively), and polynomial and exponential fits performed only marginally better than linear regressions (Figure A.7). This implies that although quartz content is important because of its effect on the restsrahlen bands, the relationships between sand content and reflectance still contains a large amount of scatter and thus uncertainty. Coefficients of determination for relationships between clay content and reflectances were less than 0.2 for all spectral bands examined, and typically much less.

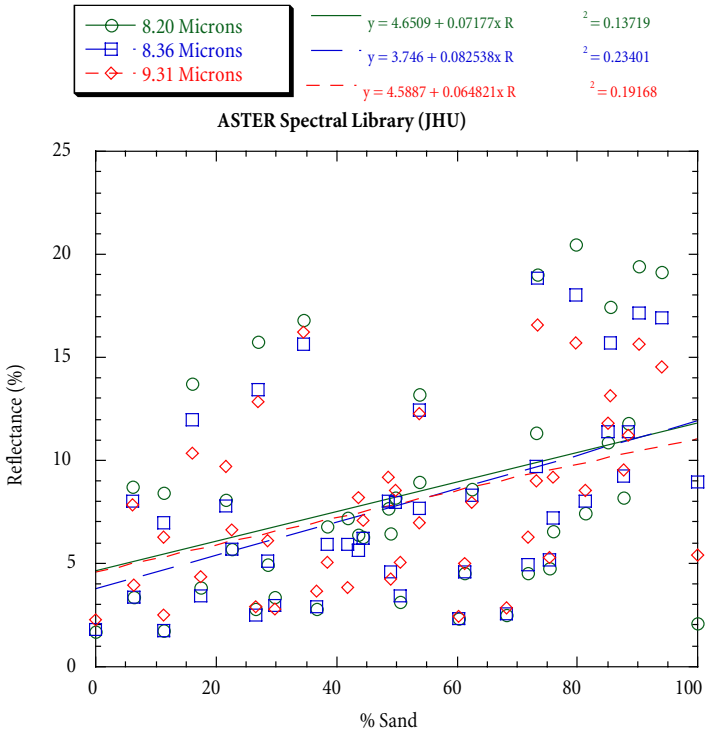


Figure A.7. Scatterplot between reflectance and the sand content of 43 soils in the ASTER spectral library for three selected spectral bands.

Moderate negative linear relationships were found between organic matter content and the reflectances at most bands in the visible/infrared and shortwave thermal infrared, with maximum values at 2.13 µm (r^2=0.4) and 3.75 µm (r^2=0.42). For these equations samples 1 and 41 were not considered. These results hint that these bands could be used to discriminate organic rich from non-organic rich soils or at least to screen those soils in order to improve the relationships to the other soil components. This would be of significance for the classification of the IPO peat class but would certainly need to be investigated further.

We also examined the usage of thermal band ratios for soil texture discrimination, as suggested by Salisbury and D'Aria (1992b). A ratio of ASTER bands 10 to 14 and 8.55 µm to 11 µm showed some improvement over the relationships found with only the 8.55 µm band. After eliminating soils with high organic matter and iron contents with a ratio of ASTER bands 10 to 12, a logarithmic fit to the data showed only moderate improvement, with sand content explaining nearly 60 percent of the variation of ASTER 10/14 (Figure A.8). Our data set and specification for ASTER band 14 is slightly different than that used by Salisbury and D'Aria

(1992b) yet our results are comparable with theirs, although they obtain an r^2 of 0.74 in their analysis.

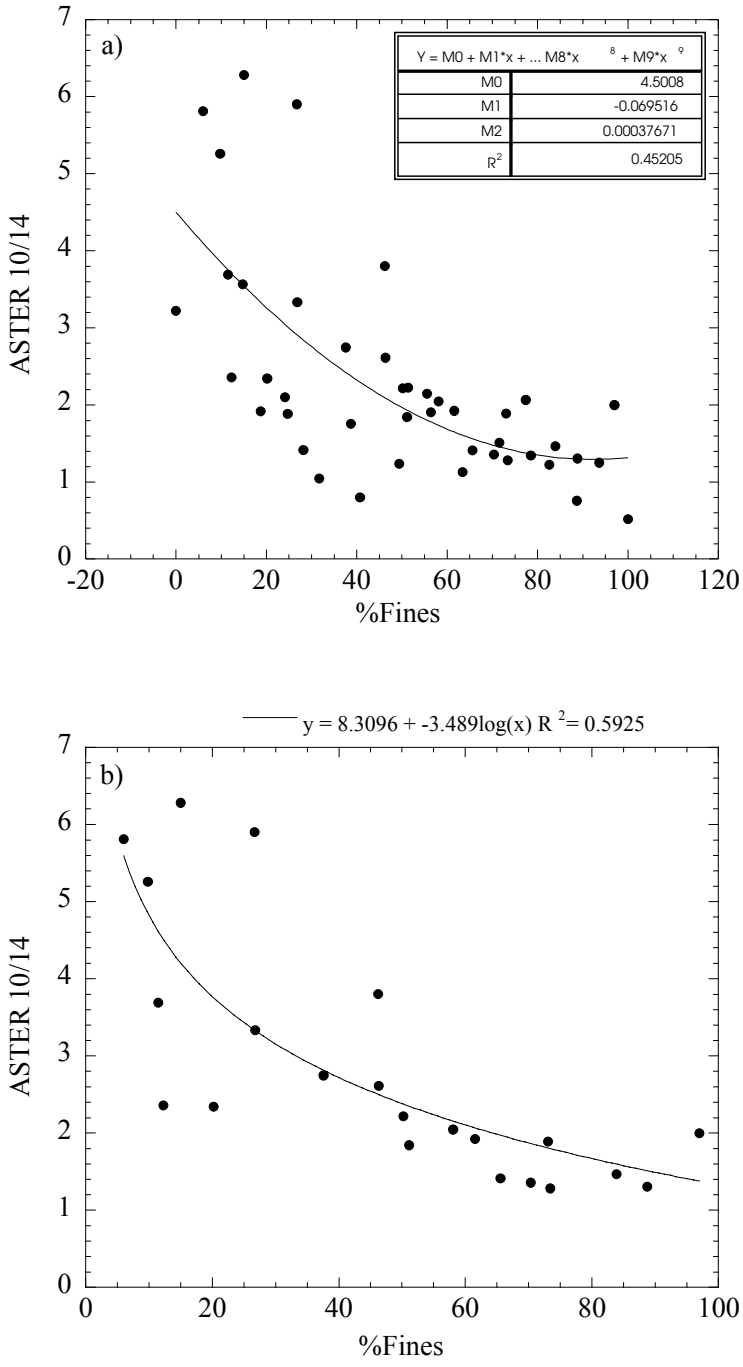


Figure A.8. a) Scatterplot showing the relationship of a ratio of ASTER bands 10 to 14 to the combined clay and silt content of all soils (Polynomial fit shown). b) Same as a) but soils with a ratio of ASTER bands 10 to 12 less than one have been screened. A logarithmic fit to the data is shown.

A.3.4 Decision Tree Analysis

Finally, we have investigated the performance of our decision tree software in separating these three broad soil categories. These were input into the C5.0 program (Quinlan 1993) with all the VIIRS spectral bands and all samples were used for training. When using all samples for training, C5.0 produced a very simple decision tree with four terminal nodes and with an overall error of 4.7 percent. Figure A.9 below shows the output of the decision tree. It is interesting to note that only VIIRS bands 9 (2.13 μm), 11 (8.55 μm), and 12 (10.5 μm) are used to partition this data set. The two soils that are misclassified by this decision tree are samples 11 and 43 which are quite different in reflectance characteristics from the other sandy or clayey soils (e.g. Figure A.6).

It should be noted that all samples were used for training the tree so that the very good results obtained may be misleading. We also performed an analysis where 50 percent of the samples were used for training and generating the tree and the other half for testing. Because of the small number of samples we performed 10 separate tests using randomly selected samples for each case. For this analysis the mean classification error for the training data was 8.2 percent (4.7 percent standard deviation), and 29.5 percent (9.5 percent standard deviation) on the testing data. These values are surprisingly good considering our previous regression analyses and are quite close to our threshold recognition accuracy. The number of soils in this analysis, however, is very small and thus the results presented here must be interpreted with caution.

A similar type of analysis was performed on simulated AVHRR reflectances from the ASTER soils. It was hoped that results from this analysis could form the basis for future work on actual AVHRR imagery. Figure A.10 shows the decision tree generated from all samples. The errors (7.0 percent) are slightly higher than those obtained for VIIRS and again the decisions utilize only the thermal bands. Results from 10 different tests with 50 percent samples were also similar to those for the VIIRS bands although with a higher mean misclassification error (10.9 percent on training, 29.5 percent on testing data, with standard deviations of 4.4 and 6.67 percent, respectively). These results are again very encouraging but will need to be verified.

The usage of thermal band ratios was also investigated. VIIRS band 11 to 13 and ASTER bands 10 to 14 ratios were included in two separate analyses. The decision tree used neither ratio when all cases were used for training the tree. However, six out of ten decision trees using the 50 percent sampling tests did use the VIIRS ratio, suggesting that this metric may be useful in addition to the individual spectral bands. Only two out of ten trees used the ASTER ratio.

The final test of decision trees involved testing the usefulness of including two thermal bands such as ASTER 10 and 12, as opposed to only one at 8.55 μm. The decision trees produced with the ASTER thermal bands did not provide any accuracy improvements over the ones with the VIIRS band, and the trees produced used ASTER bands 10 and 12 only twice each and never together. This analysis suggests that the VIIRS 8.55 μm band is sufficient to separate sandy, loamy, and clayey soils. The small sample size, however, truly limits the generalization of these analyses to broader samples, and inter-comparisons such as those performed above may need to include more representative soil types.

```
Read 43 cases from viirs.data

Decision tree:

b12 <= 2.2072: Clay (3.0)
b12 > 2.2072:
:...b11 <= 8.1915: Loam (27.0)
    b11 > 8.1915:
    :...b12 > 5.5663: Loam (4.0/1.0)
        b12 <= 5.5663:
        :...b9 <= 44.786: Sand (5.0)
            b9 > 44.786: Loam (4.0/1.0)

Evaluation on training data (43 cases):

          Decision Tree
        ----------------
        Size      Errors
           5    2( 4.7%)    <<

         (a)   (b)   (c)      <-classified as
        ----  ----  ----
           5     1            (a): class Sand
                33            (b): class Loam
                 1     3      (c): class Clay
```

Figure A.9. Output of C5.0 decision tree for 12 VIIRS bands.

```
Read 43 cases from avhrr.data

Decision tree:

avhrr4 <= 2.220898: Clay (3.0)
avhrr4 > 2.220898:
:...avhrr4 <= 3.479473: Loam (25.0/1.0)
    avhrr4 > 3.479473:
    :...avhrr5 <= 2.43947: Sand (4.0)
        avhrr5 > 2.43947: Loam (11.0/2.0)

Evaluation on training data (43 cases):

          Decision Tree
        ----------------
        Size      Errors
           4    3( 7.0%)    <<

         (a)   (b)   (c)      <-classified as
        ----  ----  ----
           4     2            (a): class Sand
                33            (b): class Loam
                 1     3      (c): class Clay
```

Figure A.10. Decision tree produced by using 5 AVHRR bands.

A.4 Soil Class Retrieval - Conclusions

The results from these analyses are mixed. On the one hand, we find large unexplained soil reflectance variability at all bands analyzed yet the decision tree appears to perform relatively well. However, the number of samples analyzed is small and may not be representative of the thousands of soil types over the Earth. Additionally, the results produced here may not be applicable when using satellite reflectances because of a variety of non-trivial problems, including atmospheric absorption and correction, surface temperature, soil moisture, and the presence of particle coatings or desert varnish (Irons *et al*., 1989; Rivard *et al*., 1992; Kealy and Hook, 1993).

The methods for soil type separation used here can only be characterized as areas of research with unproven and practically untested methodologies. Moreover, we have shown only results for three out of the six soil types and similar problems as those encountered here are anticipated for the other three soil types. Soil texture will not generally change very much over a short time scale. Is it justified to use significant resources to further develop methods to classify bare soil types on an operational basis for what really seems to be a relatively small gain?